Rising Earnings Inequality and Optimal Social Security[⋆,⋆⋆]

Pavel Brendler

University of Bonn

March 18, 2021

Abstract

Cross-sectional earnings inequality has risen sharply since the late 1970s in the United States. It remains an open question how this development has affected the insurance and income redistribution roles played by Social Security. The paper's first question is: How have the government preferences over insurance and redistribution evolved during the past four decades? I answer this question quantitatively by constructing a rich overlapping generations model. My findings indicate that the government has become less willing to provide insurance to young workers and redistribute from high-ability to low-ability households. Simultaneously, it has become more willing to tolerate income redistribution from workers toward retired households. To quantify the welfare consequences of the shift in government preferences, I ask the second question: How should Social Security have responded to inequality had the government preferences remained unchanged? Compared to the optimal policy, the current system induces a welfare loss equal to 1.2 percent in consumption equivalent terms. Finally, I show that each driving force of cross-sectional earnings inequality has a differential impact on the optimal policy.

Keywords: Optimal tax, Public pension program, Social Security reform, Earnings inequality, Idiosyncratic risk

JEL: D3, E6, H2, H3

[⋆]Institute for macroeconomics, Adenaueralle 24-42, 53113 Bonn, Germany.

[⋆⋆]I am particularly thankful for insightful and inspiring discussions with Pawel Doligalski, Thomas Hintermaier, Sagiri Kitao, Winfried Koeniger, Moritz Kuhn, Benjamin Larin, Christian Moser, and Dominik Sachs. I also thank Adrien Auclert, Christian Bayer, Yongsung Chang, Fatih Guvenen, Joachim Jungherr, Sebastian Koehne, Dirk Krueger, Patrick Macnamara, Alisdair McKay, Pontus Rendahl, Aleh Tsyvinski, Arnau Valladares-Esteban, Trine Vattø, Matt Weinzierl as well as the seminar participants at several seminars and conferences. I thank Oliver Freyermuth at the Physics Department of the University of Bonn for his excellent IT support. As always, all mistakes are mine.

Email address: `pavel.brendler@uni-bonn.de` (Pavel Brendler)

1. Introduction

Cross-sectional earnings inequality has risen sharply since the late 1970s in the United States.[1] The optimal taxation literature has already addressed the question of how this development has affected the redistributive nature of the *overall* tax and transfer system in the United States.[2] However, less attention has been devoted to the inequality's impact on distinct redistributive programs. This paper focuses on Social Security, the publicly provided pension system. Social Security is an interesting system to analyze for two reasons. First, it's a large program that plays a vital role for many U.S. households.[3] Second, Social Security is a complex system. In contrast to other programs, such as income taxation that redistributes incomes based on the individual's *current* economic conditions, the individuals' pension benefits depend on the entire *history* of their earnings.

The goal of the paper is twofold. In the first step, I study the *actual* Social Security system and ask how the government preferences over insurance and redistribution have changed since the 1970s when cross-sectional earnings inequality started to rise. I address this issue quantitatively in a rich general equilibrium overlapping generations model. The main finding is that the U.S. government has become more willing to tolerate redistribution from young workers toward retired households. Moreover, it has become less willing to insure young workers against uncertainty in their future earnings and redistribute incomes from high-ability to low-ability households. Next, I ask the second closely related question. Suppose the government tastes for insurance and redistribution have remained unchanged since the 1970s. How *should* the system have responded to inequality? Compared to the optimal policy, the existing system induces a welfare loss equal to 1.2 percent in consumption equivalent terms. I show that each driving source of cross-sectional earnings inequality has a differential impact on the effectiveness of the Social Security instruments and, therefore, the optimal policy.

Since one of the paper's goals is to understand the actual Social Security policy, the model's central ingredient is the statutory replacement rate schedule.[4] This schedule de-

[1] Heathcote et al. (2010a) and Heathcote et al. (2010b) provide extensive empirical evidence on the trends in income inequality in the United States.

[2] Among others, Lockwood and Weinzierl (2016), Heathcote et al. (2020), and Wu (2020) have made significant contributions in this field.

[3] Social Security's payouts amount to 30 percent of total government outlays. According to Hosseini and Shourideh (2019), Social Security benefits constitute as much as 40 percent of older people's total income.

[4] The replacement rate schedule is the central piece in the Social Security legislation since the Social Security Amendments of 1977 were adopted. The Amendments introduced a long-term strategy to stabilize future replacement rates and give workers clarity about their future retirement benefits. These provisions have remained essentially unchanged since 1977, but there have been several attempts to overrule them since then.

termines the individual's pension benefit based on their *average lifetime earnings*, which is essentially the average over the worker's annual earnings below the maximum taxable earnings threshold. I approximate the replacement rate schedule through a flexible function with two critical variables. The first variable sets the replacement rate of a worker, whose average lifetime earnings are equal to the economy-wide average taxable earnings. When the average replacement rate increases, the entire replacement rate schedule shifts upward, raising every current and future retiree's pension benefit. This does not occur without a cost because workers will have to pay higher payroll contributions. Hence, this policy variable generates an inter-generational conflict between the current workers and retirees. The second variable in the schedule controls its slope, i.e., the pension system's statutory progressivity. When the government reduces progressivity, all retirees whose average lifetime earnings are below the economy-wide average taxable earnings receive a lower replacement rate. By contrast, replacement rates rise for all other retirees. Thus, the second variable generates a distributional conflict between agents with high and low average lifetime earnings.

Each of the two variables in the replacement rate schedule plays a distinct role in the government's equity-efficiency tradeoff. The government can effectively use the average replacement rate to control the efficiency losses associated with the public provision of pension benefits. Simultaneously, the schedule's slope allows the government to provide insurance against labor productivity risk and achieve a less unequal distribution of resources among retired agents. However, when choosing progressivity, the government can only condition the replacement rate on the retired agent's average lifetime earnings. Therefore, the effectiveness of the second instrument will depend on the correlation between the retired agents' average lifetime earnings and their incomes. If both are highly correlated, the government can effectively leverage the pension system's progressivity to target disadvantaged groups. By contrast, the schedule's slope becomes a less effective instrument if the correlation is weak. In this case, the government has to resort to the first instrument and raise every retiree's pension benefit at the cost of higher Social Security taxes.

The government chooses the replacement rate schedule optimally by maximizing the *weighted* welfare of all agents who are alive.[5] The Pareto weights in the government's optimization problem add the necessary flexibility to help the model match the actual policy. The government sets the schedule once-and-for-all, internalizing how it will affect households along the economy's transition. The new schedule applies to all agents in the economy, in-

[5]In the real world, governments seek reelection and, therefore, propose policies to gain the support of *current* voters.

cluding those who have already retired. Given the schedule, the Social Security tax rate adjusts in each period to satisfy the government budget constraint.[6]

I calibrate two sets of parameters. The first set reflects the U.S. economy in 1979, characterized by relatively low cross-sectional earnings inequality. The second set describes the U.S. economy in 2017.[7] It reveals the substantially higher dispersion in earnings. Besides inequality, I account for population aging reflecting a divergent trend in mortality rates by ability.[8] Aging is important in the model's context because it shifts social marginal welfare weights toward older high-ability agents. Moreover, I control for changes in other aspects of the overall tax and transfer system that might affect the government's equity-efficiency tradeoff: income, capital, Medicare taxation, normal retirement age, and maximum taxable earnings threshold. Finally, I adjust the parameters that govern household preferences and the representative firm's production technology.

To address the paper's first question of how the government's preferences over redistribution have changed over time, I identify Pareto weights consistent with the actual Social Security policy in 1979 and 2017. I uncover a major disagreement *between* generations over the average replacement rate in the calibrated model economy. At the same time, there is large heterogeneity in welfare between high-ability and low-ability agents *within* each age group over pension system progressivity. To exploit this heterogeneity, I specify the Pareto weights as a function of the agent's age and ability and identify two crucial parameters in the Pareto weight function. I find that conditional on age, the relative Pareto weight on high-ability agents almost doubles over time. This findings indicates that the U.S. government has become less willing to provide insurance to young workers and redistribute incomes from high-ability to low-ability households. Simultaneously, Pareto weights shift toward older agents, conditional on their ability. More specifically, while the government in 1979 was willing to trade off one util of a 25-year-old household against two utils of the 65-year-old household of the same ability, this tradeoff is reversed in 2017. This finding suggests that the U.S. government has become more willing to tolerate redistribution from young workers toward retired households.

Although the model is salient regarding the underlying forces behind the identified shifts in Pareto weights, I provide empirical evidence that corroborates the model predictions. I

[6] I allow Social Security to be unbalanced for the model to be consistent with the payroll tax rate in the data.

[7] I choose 1979 and 2017 as these are the earliest and latest periods, respectively, for which the harmonized data on income and earnings inequality are available.

[8] In the model, the drivers of cross-sectional earnings inequality and population aging are exogenous.

specifically analyze voter turnout rates by household age and education in Congressional elections between 1978 and 2018. I show that the turnout rate of college graduates *relative* to high-school graduates of the same age has increased over time. The same observation holds along the age dimension, conditional on the household's education.

The paper's second closely related question is how Social Security *should* have responded to the economic changes had the government maintained the same preferences over insurance and redistribution as in 1979. The answer to this question allows me to quantify the welfare consequences of the shift in government preferences. In response to a combined change in all model parameters, the government optimally chooses to reduce the average replacement rate by 29 percent and increase progressivity by 2.8 times compared to 1979. This policy reduces the Social Security tax rate by 1.8 percentage points in the long run compared to 1979. The optimal policy leads to a welfare gain to the currently lived U.S. households equivalent to a 1.2 percent consumption increase. Qualitatively, the marginal effect of earnings inequality on the optimal policy turns out to be the same as the combined effect of all model parameters. Quantitatively, earnings inequality explains roughly 40 percent of the total decline in the average replacement rate and one-third of the total increase in progressivity.

Finally, I conduct a set of decomposition exercises to quantify the impact of each driving force of earnings inequality on the optimal policy. Rising college premium strengthens the correlation between retired agents' incomes and their average lifetime earnings. Therefore, the government can effectively use pension system progressivity to redistribute incomes toward income-poor retirees. The optimal progressivity increases by 18 percent compared to 1979. Since pension benefits target precisely those who need them most in this case, the government can reduce the average replacement rate by 11 percent to dampen the distortionary effect of labor taxation. Similar to the college premium experiment, higher idiosyncratic labor productivity risk leads to a substantial passthrough of cross-sectional earnings inequality into income inequality during retirement. Contrary to the previous experiment, however, there is a weaker correlation between retired agents' incomes and their average lifetime earnings. Even though the government increases progressivity by 24 percent to provide insurance to workers, it optimally chooses to uphold roughly the same average replacement rate as in 1979. This choice allows the government to redistribute resources to those workers who have relatively high average lifetime earnings but failed to accumulate sufficient savings for retirement.

Related literature

This paper relates to three strands in literature. The first strand applies the inverse optimum approach to recover social preferences for redistribution. Several studies have inverted government preferences by looking at the progressivity of the tax and transfer policy in the United States. This literature's appealing feature is that the notion of the tax and transfer system comprises a very broad range of redistributive programs, such as income taxation, social security, Medicare, child support, etc. Moreover, the tax and transfer system is specified powerfully as a function of two parameters only (the tax level and progressivity). Tsujiyama and Heathcote (2015) and Chang et al. (2018) study the progressivity of the tax and transfer system in the cross-sectional U.S. data and identify a relatively high Pareto weight attached to more productive agents. Wu (2020) and Heathcote et al. 2020 analyze the *time trends* in social preferences for redistribution. Wu (2020) estimates that the progressivity of the tax and transfer system declined during 1978–2016.[9] He constructs a rich quantitative model according to which structural economic changes, including earnings inequality and population aging, explain more than a half of the total decline in progressivity. Increased Pareto weights on high-ability households rationalize the remaining portion of the drop. Heathcote et al. 2020 challenge this finding. Their empirical investigation concludes that progressivity has remained constant between 1980 and 2016. Moreover, a *utilitarian* government in their model optimally chooses the *actual* policy in response to earnings inequality. The authors emphasize the endogenous skill investment as the critical ingredient required to match the data.

The broad view on redistribution taken by these studies poses two critical challenges. Through the aggregation of various redistributive programs, it becomes impossible to understand the exact mechanisms through which each of them affects the optimal policy. Furthermore, the mapping between the model's predictions and policy implications is everything but trivial. By focusing on Social Security, however, I can closely analyze its distinct role in providing insurance and redistributing incomes between and within generations. I identify the evolution of government's preferences over inter-generational and intra-generational income redistribution. One cannot disentangle these motives in the existing literature. Moreover, my model offers predictions that have straightforward policy implications.

Lockwood and Weinzierl (2016) study the statutory *income tax* policy and find a significant increase in the average marginal social welfare weights for high-income households

[9] Note that his progressivity measure excludes Social Security benefits, as he focuses on the working age population.

during 1980–1990.[10] In a closely related work that also focuses on Social Security (Brendler, 2020), I analyzed the *inter*-generational preferences for redistribution and inferred a trend in age-dependent Pareto weights. One limiting feature of that study is that the government was restricted to control the inter-generational income redistribution only. Hence, the paper was unable to study the impact of rising earnings inequality on insurance provision and *intra*-generational income redistribution. The current paper emphasizes the importance of the *joint* distribution of Pareto weights by age and ability for the optimal policy because the agents' most preferred policies vary along these two dimensions.

The second strand in the literature asks a *normative* question of how a public pension system should optimally look. Hosseini and Shourideh (2019) study Pareto optimal policy reforms with heterogeneous mortality rates and time preferences. Ndiaye (2020) examines lifecycle taxation with endogenous retirement. Moser and Olea de Souza e Silva (2019) construct a model in which Social Security and income taxation arise as the decentralization of an optimal policy that trades off savings adequacy (due to present bias heterogeneity) with income redistribution (due to ability heterogeneity).[11] As opposed to my work, these studies primarily focus on the decentralization of the first-best policies. In a related study, Huggett and Parra (2010) conduct a Social Security reform by optimally choosing the parameters of the existing pension benefit function. While they find a small welfare gain in the model's version without idiosyncratic labor productivity risk, the welfare gain becomes substantial once they add persistent and temporary earnings shocks into the model. The novel feature of my work is to analyze the implications of the rising trend in cross-sectional earnings inequality on the optimal Social Security reform.[12]

The third strand of related work has analyzed the macroeconomic and welfare consequences of different retirement financing reforms. Conesa and Krueger (1999), Huggett and Ventura (1999), Nishiyama and Smetters (2007), Fuster et al. (2007), Kitao (2014), McGrattan and Prescott (2017), and Nishiyama (2019) have made significant progress in this field.

[10] Chang et al. (2021) provide an interesting insight that the discrepancy between the actual and the utilitarian income tax policy reduces once one accounts for the ex-ante heterogeneity in workers' ability and income-dependent voter turnout rates.

[11] Contrary to Tsujiyama and Heathcote (2015), the authors recover a hump-shaped distribution of Pareto weights that puts relatively more weight on the second ability quartile mostly because they match the retirement savings system (which is fairly progressive) jointly with the income tax system (which is fairly regressive). Qualitatively, this finding is similar to Jacobs et al. (2017), who apply the revealed preference approach to the Netherlands' income tax policy.

[12] There is a set of other important studies. Fehr and Habermann (2008) and Fehr et al. (2013) analyze the optimal progressivity of the German pension system and show that progressivity matters quantitatively for households' welfare.

This literature has studied exogenous and arguably politically infeasible reforms (e.g., complete elimination of Social Security). By contrast, my paper rationalizes the existing Social Security system. As argued by Lockwood and Weinzierl (2016) and Stantcheva (2016), the distribution of Pareto weights captures feasibility constraints imposed on the political process. Hence, the weights can be applied in policy analyses to shrink the set of all economically feasible proposals to those that are also implementable from the political standpoint.

2. Model

2.1. Demographics

The economy is populated by overlapping generations of agents. Each period a continuum of agents is born. The birth rate equals n. Age is denoted by j. Agents enter the economy and start working at age $j = 1$, which corresponds to real-life age 25. The mandatory retirement age is J^R. When I calibrate the model below, this age will correspond to real-life age 65 in the baseline version of the model. Each generation lives for J periods, which is 85 years in real-life terms.

At any point in time, I normalize the total population size to 1. Let μ^W be the total mass of working-age agents such that $\mu^R = 1 - \mu^W$ is the share of retired agents. At birth, each individual receives a realization of ability, which is a random variable $z \in Z = \{H, L\}$, where H stands for high-ability and L – for low-ability. Once drawn, the ability remains constant throughout agent's life. The share of high-ability agents among the newborns is denoted by λ_H, so that $1 - \lambda_H$ is the fraction of low-ability agents. When calibrating the model, high-ability agents will correspond to household heads with at least a completed college degree in the data, while low-ability agents will correspond to all the remaining households.

Ability plays two important roles in my model. First of all, it controls agent's mortality rates over the lifecycle. Denote by $\psi_{z,j}$ the probability that an agent with ability z survives up to age $j+1$, conditional on surviving up to age j. Second, ability controls the deterministic and stochastic elements of agent's labor productivity as explained below.

2.2. Production

A representative firm produces the final output good according to the production function:

$$Y_t = K_t^{\theta} {N_t}^{1-\theta}, \tag{1}$$

where K_t – the aggregate capital stock, N_t – the aggregate effective labor input and $\theta \in (0, 1)$ – the capital share in production. I explicitly show the dependence of model variance on time using the index t, since the transitional dynamics will be an essential part of the model description.[13] The output can be consumed or invested in capital. The depreciation rate of capital is $\delta \in (0, 1)$. The firm produces output goods and sells them in a competitive market at a price that is normalized to one. The rental price of capital, r_t, and the wage per effective unit of labor, w_t, are determined competitively:

$$r_t = \theta \left(K_t/N_t\right)^{\theta-1} - \delta \text{ and } w_t = (1-\theta)\left(K_t/N_t\right)^{\theta}. \tag{2}$$

2.3. Households

2.3.1. Overview

Agents are born with zero assets but can accumulate savings over time. Households hold two types of assets: shares in the representative firm and government bonds. Both assets bear no risk and generate a pre-tax rate of return r_t defined in (2). Since households are indifferent between investing in government bonds or firm shares, I denote the quantity of either asset held by an agent in the current period by a and the amount saved for the next period by a'.[14] Borrowing is ruled out, i.e. $a' \geqslant 0$.

Agents must pay a proportional capital tax τ_a on the asset income. By the no arbitrage condition, both assets must be paying the same after-tax rate of return which I label by:

$$\tilde{r}_t = (1 - \tau_a) r_t. \tag{3}$$

Finally, all agents in the economy pay a proportional tax τ_c on consumption c.

2.3.2. Worker's labor productivity

A worker is an agent of age $j = 1, ..., J^R - 1$. Each worker is endowed with one unit of productive time in each period, a fraction of which she supplies optimally to a competitive labor market. Agent's productivity is composed of a deterministic and a stochastic component. Each of these two components will be essential in the experiment section of the paper, when I model the rise in earnings inequality since the late 1970s.

The deterministic element of agent's productivity is denoted by $\zeta_{z,j}$. I will calibrate $\zeta_{z,j}$ to match the wage profiles of each ability type in the data. Everything else equal, high-ability agents receive a wage premium, $\zeta_{H,j}/\zeta_{L,j}$, over low-ability agents.

[13] In the next section, I will calibrate the model to two distinct steady states. Some of the parameters, such as the share of high ability agents λ_H, will take on different values in each of the steady states. To simplify notation, I avoid the time index for these parameters.

[14] Throughout the paper, I omit the time index for individual variables to simplify notation.

The stochastic element of agent's productivity is denoted by $y_{i,z,j,t}$. It consists of a persistent auto-regressive shock η and a transitory shock ν:

$$y_{i,z,j,t} = \eta_{i,z,j,t} + \nu_{i,z,t}, \tag{4}$$

The persistent shock follows an AR(1) process:

$$\eta_{i,z,j,t} = \rho_z \eta_{i,z,j-1,t-1} + \gamma_{i,z,t} \text{ with } \eta_{i,z,1,t} = 0, \tag{5}$$

where ρ_z is a constant persistence parameter and the error terms are distributed as follows: $\nu_{i,z,t} \sim \mathcal{N}(0, \sigma^2_{\nu,z})$, $\gamma_{i,z,t} \sim \mathcal{N}(0, \sigma^2_{\gamma,z})$. The conditional variance of $\eta_{i,z,j,t}$ increases with age according to:

$$\mathsf{Var}(\eta_{i,z,j,t}) = \sigma^2_{\gamma,z} \times \sum_{h=0}^{j-1} \rho_z^{2h} \text{ for } j \geqslant 1. \tag{6}$$

For $|\rho_z| < 1$, the expression above converges to $\sigma^2_{\gamma,z}/(1-\rho_z^2)$ which is the unconditional variance of the AR(1) process in (5).

Observe that each ability group shares separate stochastic processes, which will allow me in the calibration section to capture any potential effects of education on the stochastic process for earnings.

To simplify notation below, I drop the individual index i and the time index t and stack the realizations of $\eta_{z,j}$ and ν_z into a vector $\mathbf{y} = \mathbf{y}_{z,j} \in \mathcal{Y}$. The stochastic process for $\mathbf{y}$ follows a finite-state Markov process with stationary transitions over time:

$$\pi(\mathbf{y}, \mathcal{Y}) = \mathsf{Prob}(\mathbf{y}_{z,j+1} \in \mathcal{Y} \mid \mathbf{y}_{z,j} = \mathbf{y}). \tag{7}$$

Let $\Pi_{\mathbf{y}}$ denote the invariant probability measure of newborn agents with productivity $\mathbf{y}$.

Summarizing, the total labor productivity (per unit of raw labor) of a worker with ability z and of age j is given by $\zeta_{z,j} \times \exp(y_{z,j})$. To simplify notation, I denote agent's total labor productivity by ϵ.

2.3.3. Worker's budget constraint

A worker supplies raw labor $l \in [0,1]$ to the competitive labor market and receives *gross earnings* equal to:

$$e = w_t \epsilon l,$$

where w_t is the wage rate per unit of effective labor defined in (2). Earnings are subject to Social Security taxation. Define worker's earnings taxable for the Social Security purpose as:

$$\tilde{e}_{SS} = \min(cap_{SS,t}, e),$$

where $cap_{SS,t}$ is the maximum taxable earnings threshold above which earnings remain untaxed. The Social Security tax burden borne by a worker is then given by $\tau_{SS,t}\tilde{e}_{SS}$, where $\tau_{SS,t}$ is a linear Social Security tax rate.[15] As will become evident below, the government sets the pension benefits, while $\tau_{SS,t}$ adjusts to balance the government budget constraint. This explains the time index in $\tau_{SS,t}$.

During working career, an agent accumulates average lifetime earnings, $\bar{e}$, that determine her pension benefit during retirement, as will be explained below. The law of motion for average lifetime earnings reads as follows:

$$\bar{e}' = \begin{cases} \left[(j-1) \times \bar{e} \times \tilde{E}_t/\tilde{E}_{t-1} + \tilde{e}_{SS}\right]/j & \text{if } j < J^R \\ \bar{e} & \text{if } j \geqslant J^R \end{cases}. \tag{8}$$

where $\bar{e}'$ denotes the agent's next period average lifetime earnings. The first line of the equation applies to workers. It reflects the still active provision of the Social Security Amendments of 1977 according to which the individual's average lifetime earnings are indexed to the economy-wide average taxable earnings, $\tilde{E}_t$, one period prior to the agent's retirement. All workers enter the labor market with no prior earnings histories, i.e. $\bar{e} = 0$ for $j = 1$. The second line states that the individual's average lifetime earnings remain constant during retirement.

Apart from Social Security contributions, all workers also pay a linear tax, τ_M, to finance the Medicare's hospital insurance program. Using similar notation as for Social Security taxes, define worker's Medicare-taxable earnings as:

$$\tilde{e}_M = \min(cap_{M,t}, e),$$

where no Medicare taxes are imposed on earnings above $cap_{M,t}$. In accordance with the Social Security legislation, the taxable maximum in the Social Security and Medicare programs are automatically adjusted to the growth in the average taxable earnings.[16]

[15] I abstract away from the disability insurance and therefore ignore disability taxes.

[16] See Appendix A for details.

Furthermore, all workers pay income taxes according to an income tax function Λ that depends on agent's pre-tax earnings.[17]

Putting all the ingredients together, the worker's budget constraint reads:

$$a' + (1+\tau_c)c = (1+\tilde{r}_t)a + e - \tau_{SS,t}\tilde{e}_{SS} - \tau_M \tilde{e}_M - \Lambda(e). \tag{9}$$

2.3.4. Retired agents

At age J^R, all agents enter compulsory retirement. During retirement, agents receive a pension benefit B_t. I will explain the pension benefit function in more detail below. For now, it is sufficient to note that it depends on a two dimensional vector of Social Security policy variables $\boldsymbol{\alpha}_t$.

Apart from the pension benefit, retirees receive a lump-sum Medicare transfer denoted by $T_{M,t}$. Similar to workers, retired agents must pay capital taxes and consumption taxes, which I described above.[18]

Summarizing, the budget constraint of a retired agent reads:

$$a' + (1+\tau_c)c = (1+\tilde{r}_t)a + B_t + T_{M,t}. \tag{10}$$

2.3.5. Agent's optimization problem

Let $\mathbf{x}$ and $\mathbf{x}'$ denote the individual's current and future states, respectively:

$$\mathbf{x} = (z, j, \mathbf{y}, a, \bar{e}) \text{ and } \mathbf{x}' = (z, j+1, \mathbf{y}', a', \bar{e}').$$

Recall that by assumption agent's ability z remains constant throughout agent's life. The laws of motion for $\mathbf{y}'$ and $\bar{e}'$ are given by (7) and (8), respectively, while a' is pinned down by the budget constraint for workers in (9) and for retirees in (10). Furthermore, let $F_t(\mathbf{x})$ be the cumulative population density function of agents over state $\mathbf{x}$ at time t and $f_t(\mathbf{x})$ – the corresponding density function.

Assume that at time t the government sets a constant future Social Security policy, $\boldsymbol{\alpha}$, that becomes effective the following period (more details follow below). Taking $(\boldsymbol{\alpha}_t, \boldsymbol{\alpha})$ as

[17] I allow income taxes to depend on the pre-tax earnings e to be consistent with how the parameters of Λ are going to be estimated in the data. See section 3.

[18] I assume that pension benefits are not taxed.

given, agents solve the following dynamic programming problem at time t:

$$\begin{aligned} V(x;\boldsymbol{\alpha}_t,\boldsymbol{\alpha},F_t) &= \max_{c,l,a'} u(c,l)+\beta\times\psi_{z,j}\times\mathbb{E}\left[\tilde{V}(x';\boldsymbol{\alpha},F_{t+1})\mid(x,t)\right] &(11)\\ &+ (1-\psi_{z,j})\times\phi(a')+\mathbb{1}_{j\geqslant 35}\times\chi\times\bar{u}_{j-34,t} &(12)\end{aligned}$$

subject to the budget constraint for the working-age agent (9) and for the retired agent (10). The solution to this problem generates the decision rules for consumption, labor and savings, which I denote briefly by $c^\star$, $l^\star$ and $a'^\star$, respectively. Worker's pre-tax earnings are then given by $e^\star = w_t \epsilon l^\star$.

In (11), u is agent's instantaneous utility function, β – the discount factor, and $\mathbb{E}$ – a conditional expectation operator. V denotes the discounted lifetime indirect utility of agent in state x at time t when the current policy is $\boldsymbol{\alpha}_t$ and the future(constant) policy is $\boldsymbol{\alpha}$. The value function on the right-hand side, $\tilde{V}_t$, represents the agents' welfare associated with the permanent policy $\boldsymbol{\alpha}$.[19]

If an agent deceases at time t, she receives an instantaneous "warm-glow" utility from bequeathing her asset holdings denoted by $\phi(a')$ in (11). This feature of the model will help me better fit inequality at the upper tail of the wealth distribution in the data.

Even though I do not model families, I introduce a simple dynastic link between generations. Particularly, *parents* of model-age 35 and above (therefore, an indicator function in 11) receive additional utility from the average (instantaneous) utility of their *children* who are assumed to be 34 years younger. The instantaneous utility of a child whose parent is j years old at time t is given by $\bar{u}_{j-34,t}$.[20] Parameter χ governs the extent to which *parents* care about their *kids*.

2.4. Government

The government runs three activities, each having its separate budget. First of all, the government operates the Social Security program by collecting payroll contributions from workers and paying pensions to the retirees. I allow the government to accumulate deficit

[19] The value function $\tilde{V}$ is defined as follows:

$$\begin{aligned} \tilde{V}(x;\boldsymbol{\alpha},F_{t+1}) &= \max_{c,l,a'} u(c,l)+\beta\times\psi_{z,j}\times\mathbb{E}\left[\tilde{V}(x';\boldsymbol{\alpha},F_{t+2})\mid(x,t+1)\right]\\ &+ (1-\psi_{z,j})\times\phi(a')+\mathbb{1}_{j\geqslant 35}\times\chi\times\bar{u}_{j-34,t+1}.\end{aligned}$$

Thus, starting from period $t+1$, the problem becomes recursive.

[20] See Appendix A for the definition of $\bar{u}_{j-34,t}$.

in the Social Security system by issuing government bonds to households. Second, the government collects consumption and capital taxes from all agents, pools them together with the wealth left by deceased agents and net income taxes. It uses these resources to pay for an exogenous government spending. Third, the government administrates a balanced budget Medicare program by collecting hospital insurance taxes from workers and paying a lump-sum Medicare transfer to the retirees.

Below I describe each of these three activities in detail.

2.4.1. Social Security

Replacement rate schedule

One of the central ingredients in the model is the replacement rate schedule that determines the agent's pension benefit upon reaching the mandatory retirement age. I restrict the replacement rate schedule to the class of policies defined by the function[21]:

$$R(\hat{e}; \boldsymbol{\alpha}_t) = \alpha_{1,t} \times [\alpha_{2,t} + (1 - \alpha_{2,t}) \times \hat{e}] \,, \tag{13}$$

where $\alpha_{1,t} \in \mathcal{R}_+$ and $\alpha_{2,t} \in \mathcal{R}$ are the two policy instruments introduced above, i.e. $\boldsymbol{\alpha}_t = (\alpha_{1,t}, \alpha_{2,t})$. I will elaborate on how each of these two variables affects the schedule after introducing some additional ingredients.

The replacement rate is defined as a function of the agent's average lifetime earnings *normalized* by the economy-wide Social Security taxable earnings in the last period of this agent's working career, i.e. $\hat{e} = \bar{e}/\tilde{E}_{t-j+J^R-1}$ for $j \geqslant J^R$.[22] Observe that once the agent enters retirement, her normalized lifetime earnings $\hat{e}$ remain unchanged, since $\tilde{E}_{,t-j+J^R-1}$ is pre-determined and $\bar{e}$ remains constant for $j \geqslant J^R$ according to (8). Thus, the replacement rate of a retired agent remains unchanged as long as $\boldsymbol{\alpha}_t$ remains constant.

Given the schedule of replacement rates, the agent's pension benefit reads:

$$B_t = \bar{e} \times R(\hat{e}; \boldsymbol{\alpha}_t). \tag{14}$$

---

[21] Additionally, I require the replacement rates to be non-negative.

[22] Normalizing the average lifetime earnings by the economy-wide taxable earnings is consistent with one of the key goals of the Social Security Amendments of 1977 which was to ensure that an individual's replacement rate only depends on this individual's relative earnings position averaged over her working lifetime. Two workers who enter retirement at different points in time but who have had the same growth-adjusted average lifetime earnings relative to the economy-wide taxable earnings should receive the same replacement rate. This idea is consistent with the replacement rate specification in (13) coupled with the law of motion for $\bar{e}$ in (8).

Given B_t, the Social Security tax rate, $\tau_{SS,t}$, adjusts period-by-period to satisfy the Social Security budget constraint:[23]

$$\tau_{SS,t}\mu^W\tilde{E}_t + D_{t+1} = \int_{x:j\geqslant J^R} B_t dF_t(x) + (1+\tilde{r}_t)D_t. \tag{15}$$

The left-hand side of this constraint shows the total government revenue at time t, where $\mu^W\tilde{E}_t$ are the aggregate Social Security taxable earnings and D_t stands for government debt issued in the previous period. I assume that the amount of outstanding debt in every period must satisfy a fixed debt-to-GDP ratio: $d_{SS} = D_t/Y_t$. The right-hand side represents the total government expenditure which consists of pension payments and the service of outstanding debt at a no arbitrage borrowing cost $\tilde{r}_t$ defined in (3).

Each instrument in the replacement rate schedule (13) has an economic interpretation. The variable $\alpha_{1,t}$ controls the replacement rate of an individual whose average lifetime earnings at retirement are precisely equal to the economy-wide average taxable earnings, i.e. $R(\hat{e} = 1; \boldsymbol{\alpha}_t) = \alpha_{1,t}$. For the sake of brevity, I will refer to $\alpha_{1,t}$ as the *average replacement rate* below. When $\alpha_{1,t} = 0$, the entire pension system shuts down. As the government increases $\alpha_{1,t}$, the replacement rate schedule shifts upward, raising pension annuities of all current and future retirees in the economy. Therefore, through the government's budget constraint in (15), $\alpha_{1,t}$ will have a strong quantitative impact on the Social Security contributions paid by the working-age population.[24] Hence, this policy variable generates an inter-generational conflict between the current workers and retirees.

The variable $\alpha_{2,t}$ controls the pension system's statutory progressivity. All else equal, an increase in $\alpha_{2,t}$ raises the replacement rate of all individuals whose average lifetime earnings at retirement are below the economy-wide average (i.e., those with $\hat{e} < 1$) and lowers the replacement rate for all individuals with $\hat{e} > 1$, without affecting the replacement rate at the average economy-wide earnings. As I will illustrate in Section 4.3, changes in $\alpha_{2,t}$ have only a minor quantitative effect on the equilibrium level of the Social Security tax rate, $\tau_{SS,t}$, in the calibrated model economy because a rise in total pensions accruing to one group of retirees is approximately offset by the drop in the total pension amount flowing to the other group. The major role of $\alpha_{2,t}$ is to control insurance provision and ex-post income redistribution *within* the pool of retired agents. For this reason, this variable generates an intra-generational

[23] To simplify notation, I use the integral to sum over x throughout the paper. Note, however, that age and ability are discrete variables.

[24] I will confirm this point in Section 4.1.

conflict in the model.

Government maximization problem

At time t, the government faces a given replacement rate schedule $\boldsymbol{\alpha}_t$. In the same period, the government makes an unanticipated announcement that it will implement policy $\boldsymbol{\alpha}^\star$ which becomes effective in period $t+1$ and remains constant in all subsequent periods. The new schedule affects all agents in the economy, including those who have already retired.[25] When choosing $\boldsymbol{\alpha}^\star$, the government maximizes the weighted sum of expected discounted lifetime utilities of all generations who are *currently* alive.[26] Formally, the government solves in period t:

$$\boldsymbol{\alpha}^\star = \arg\max_{\boldsymbol{\alpha}} \int_x \omega(j, z; \kappa_t) V(x; \boldsymbol{\alpha}_t, \boldsymbol{\alpha}, F_t) dF_t(x), \tag{16}$$

subject to the Social Security budget constraint in (15). Note that the distribution of agents at time t, F_t, is given and does not depend on the choice of $\boldsymbol{\alpha}$.

In the objective function above, ω is the Pareto weight function specified as follows:

$$\omega(j, z; \kappa_t) = \begin{cases} j^{\kappa_{1,t}} & \text{if } z = L \\ j^{\kappa_{1,t}} \times \kappa_{2,t} & \text{if } z = H \end{cases}, \tag{17}$$

where $\kappa_{1,t} \in \mathcal{R}$ and $\kappa_{2,t} \in \mathcal{R}_+$ are parameters. Below I refer to $\kappa_{1,t}$ as the *age bias* and $\kappa_{2,t}$ – as the *ability bias*.

When $\kappa_{1,t} = 0$ and $\kappa_{2,t} = 1$, the government's maximization problem boils down to a utilitarian social welfare function. When $\kappa_{1,t} > 0$, however, the government puts a higher weight on older agents, conditional on their ability. The opposite is true when $\kappa_{1,t} < 0$. When $\kappa_{2,t} > 1$, the government puts a larger weight on a high-ability agent compared with a low type of the same age. The opposite is the case when $0 \leqslant \kappa_{2,t} < 1$.

The government chooses the replacement rate schedule once-and-for-all. This assumption is in line with the spirit of the Social Security Amendments of 1977, which were adopted as a long-term policy intended to give current young generations clarity about their future retirement benefits. Furthermore, the government chooses a constant replacement rate schedule instead of a constant Social Security tax rate. Again, this assumption is consistent with the

[25] This assumption is consistent with the fact that the contributors have, in principle, no legal entitlement to receive a certain pension benefit in the United States.

[26] One of the paper's goals is to understand the actual policy. In the real world, governments seek reelection and propose policies to gain the support of *current* voters.

Social Security Amendments' intention to stabilize future replacement rates.

The choice of the Pareto weight function in (17) requires a justification. First of all, the chosen specification allows me to identify the parameter vector $\boldsymbol{\kappa}_t$ in the data. As I will show below, α_1 in the data identifies the age bias in the model because older agents, on average, prefer a higher level of the replacement rate schedule. At the same time, the pension system's progressivity, α_2, identifies the ability bias in the model because, conditional on age, high-ability agents prefer, on average, a less progressive pension system than low-ability agents. Second, both arguments of the Pareto weight function – age and ability – are part of the agent's state space $\boldsymbol{x}$ and, therefore, do not depend on policy $\boldsymbol{\alpha}$, which drastically simplifies the identification of Pareto weights.[27] Third, one could include other dimensions of heterogeneity among households as arguments of the Pareto weight function. This is a promising approach because it may uncover additional dimensions of heterogeneity and, therefore, distributional conflicts among the households. This approach's major obstacle is computational intensity because identification requires that the number of parameters in the weighting function is at least as large as the number of instruments available to the government. I leave this direction to future research.

2.4.2. Income transfer program

Apart from running Social Security, the government administrates a balanced budget income transfer program. More specifically, the government collects net income taxes from workers according to the income tax function Λ. Besides, the government collects capital taxes and consumption taxes from all agents in the economy and confiscates wealth left by deceased agents. These resources are pooled together and used in the same period to pay for government spending, G_t, which is wasted in the context of this model.

Summarizing, the government budget constraint in period is given by:

$$G_t = \int_{\boldsymbol{x}:j<J^R} \Lambda(e^\star) dF_t(\boldsymbol{x}) + \tau_c C_t + \tau_a r_t A_t + \Phi_t, \tag{18}$$

where C_t, A_t and Φ_t denote aggregate consumption, savings and bequests, respectively.

[27] This would not be the case if the agent's Pareto weight would be a function of her current earnings, her expected average lifetime earnings at retirement, etc.

2.4.3. Medicare

Finally, the government administrates in each period the Medicare program by collecting hospital insurance taxes from workers and paying a lump-sum transfer, $T_{M,t}$, to retirees:

$$\tau_M \times \int_{x:j<J^R} \tilde{e}_M dF_t(x) = T_{M,t} \times \mu^R. \tag{19}$$

2.5. Competitive equilibrium

Appendix A defines the competitive equilibrium of the model.

3. Calibration

3.1. Overview

I calibrate two sets of parameters. The first set reflects the U.S. economy in 1979, whereas the second describes the economy in 2017. I choose these time periods because they are the earliest and latest periods, respectively, for which the harmonized Current Population Survey (CPS) data are available (see below). I assume that the economy is in a steady state in 1979 and 2017. Below, I will refer to the model economy calibrated under either set of parameters as the *baseline* model. To simplify notation, I drop time index t throughout this section. One model period equals one year. Agents enter the model at age 1 which corresponds to a real-life age 25. The maximum possible age is $J = 61$ (real-life age 85).

Table 1 shows the parameters of the model that remain constant in both steady states. Table 2 displays the parameters that take on different values in each steady state. These parameters are divided into six subgroups: 1) Inequality, 2) Social Security, 3) Utility and technology, 4) Aging, 5) Income taxes, and 6) Other taxes. The parameters marked by an asterisk in the table were calibrated inside the model. Other parameters were either calibrated outside the model or assigned values from external sources. I will next describe the calibration strategy in detail.

3.2. Calibration strategy

Inequality

I use the harmonized time series data from the CPS to compute empirical moments of inequality.[28] An agent in the model corresponds to a household in the data. I use household-

[28]The CPS extracts harmonized across all years during 1979–2018 are publicly available at http://ceprdata.org/cps-uniform-data-extracts/march-cps-supplement/march-cps-data/.

Table 1: Constant Parameters of the Model.

Parameter	Description	Value
J	Life span	61 (real-life age 85)
σ	Risk aversion	2
χ	Degree of altruism	0.5
(ϕ_1, ϕ_2)	Bequests	$(-9.5, 11.6)$
ρ	AR(1) coefficient	0.97

Notes: The parameter values shown in the table remain constant across the steady states. The values for (J, σ, χ) are fixed, while the values for (ϕ_1, ϕ_2, ρ) are borrowed from external sources.

level data as opposed to individual-level data to account for the insurance against idiosyncratic labor productivity risk among household members, the research of which has demonstrated to be quantitatively important (see Fuster et al., 2007). My CPS sample includes both male- and female-headed households age 25–64 for 1979 and 25–65 for 2017, where the difference in the range is due to the difference in the calibrated mandatory retirement age J^R (see below). A high-ability agent corresponds to a household head with at least a completed college degree in the CPS. Otherwise, the household head is a low-ability agent. I keep only those household heads who work at least 260 hours.[29] In the model, I drop all agents who supply less than 5 percent of their unitary time endowment to be consistent with this sample selection criterion.[30]

The age-efficiency profile, $\{\zeta_{z,j}\}_{z,j=1}^{J^R-1}$, controls the deterministic portion of agent's earnings over her lifecycle. Following Hansen (1993), I construct $\zeta_{z,j}$ as follows. First, I compute mean hourly earnings by age separately for high school graduates and college graduates using the 1980 CPS extract. Second, I apply a quadratic polynomial curve to each profile to extract age-dependent variation in earnings. Finally, I normalize the fitted profiles by the average hourly earnings, computed on a pooled sample of households. I proceed similarly to the 2018 CPS extract.

I do not visualize the computed age-efficiency profiles. Instead, Figure 1 contrasts the im-

[29] My further restrictions on the sample are as follows. I drop a household if at least one of its members reports strictly positive earnings but zero hours worked. I also drop all observations with non-positive household earnings. Since a fraction of households reports earnings and hours worked, which imply a wage rate below the minimum wage rate, I drop all the households located in the bottom 1 percent of the household earnings distribution in a given year.

[30] In the Time Use Surveys of 2003–2005, as documented by Carceles-Poveda and Abraham (Unpublished), the household's disposable time is 97 hours per week after deducting sleep and personal care, which sums up to roughly 5,096 hours annually. The threshold of 5 percent is then computed as 260/5,096, which corresponds to supplying fewer than 260 hours annually.

Table 2: TIME-VARYING MODEL PARAMETERS.

Parameter	Description	1979	2017
1) Inequality:			
$\{\zeta_{z,j}\}_{z,j=1}^{J^R-1}$	Age-efficiency profile	Figure 1	
$(\sigma^2_{\gamma,H}, \sigma^2_{\gamma,L})$	Persistent shock variance, 10^{-2}	$(0.58, 0.42)$	$(1.26, 0.61)$
$(\sigma^{2\star}_{\nu,H}, \sigma^{2\star}_{\nu,L})$	Transitory shock variance, 10^{-2}	$(0.56, 0.6)$	$(0.7, 0.65)$
λ_H	Share of college graduates, %	25	43
2) Social Security:			
α_1	Statutory average replacement rate	0.45	0.5
α_2	Statutory degree of progressivity	1.69	1.48
J^R	Retirement age	41	42
$cap^\star_{SS}$	Taxable maximum	1.11	1.64
d_{SS}	Government deficit in GDP, %	0.1	0.47
3) Utility and Technology:			
$\gamma^\star$	Weight on consumption	0.46	0.44
$\beta^\star$	Discount factor	1.01	0.99
θ	Capital share	0.43	0.46
δ	Depreciation rate of capital, %	8	6
4) Aging:			
$\{\psi_{z,j}\}_{z,j=1}^{J}$	Age-profile of survival rates	Figures C.9–C.10	
n	Birth rate, %	0.62	1.3
5) Income Taxes:			
$\tau^\star_I$	Average level of taxation	0.22	0.14
$\tilde{\tau}_I$	Degree of progressivity	0.19	0.14
6) Other Taxes:			
τ_a	Capital tax, %	38.4	33
τ_c	Consumption tax, %	5.3	4.1
τ_M	Hospital insurance tax, %	2.1	2.9
cap_M	Taxable maximum in Medicare	1.11	–

Notes: The table shows the parameters that take on different values in each of the two steady states. The parameters calibrated inside the model are marked with an asterisk.

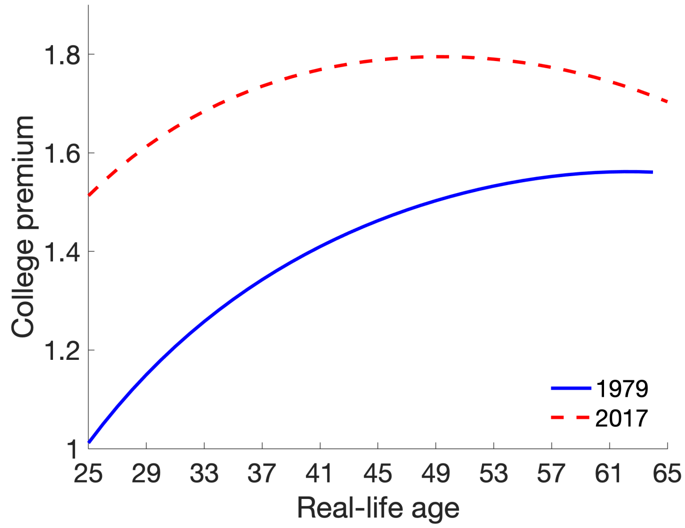

Figure 1: College Premium in the Data, by Age.

Notes: The figure displays the ratio of age-efficiency units between college graduates and high school graduates (by age) computed for 1979 and 2017 in the CPS. See main text for details.

plied profiles of college premium, defined as $\zeta_{H,j}/\zeta_{L,j}$, in each year. Two observations emerge from the figure. First, a college degree graduate received a wage premium over the high-school graduate at each age in both periods. Admittedly, this premium was almost insignificant for a 25-year-old college graduate in 1979. However, the wage differential increased steeply for other age groups, with a 63-year-old college graduate receiving almost a 50 percent higher wage. Second, the wage premium curve shifts upward in 2017. The largest increase (more than 50 percent) experienced a 25-year-old college graduate. The documented shift in the college premium will be one of the main drivers of earnings inequality in the model between 1979 and 2017.

I also account for the increased share of college graduates during 1979–2017. More specifically, the share of high-ability agents, λ_H, almost doubles from 25 percent in 1979 to 43 percent in 2017. These numbers are consistent with the rise of household heads with a completed college degree in the CPS.

Next, I estimate the idiosyncratic component of labor productivity specified in (4) using the residual variation in earnings (i.e., the variation that is left in the data after extracting the age-dependent variation). I assume that the persistence parameter, ρ, is independent of agent's ability and constant over time. I borrow $\rho = 0.973$ from Heathcote et al. (2010b), who estimate the stochastic process in (4) on a pooled PSID sample of high school and college graduates during 1967–2000.[31] Then I estimate the variances of the persistent shock, $\sigma^2_{\nu,z}$, and the variances of the temporary shock, $\sigma^2_{\gamma,z}$, outside the model by fitting the stochastic process in (6) to the empirical profiles of earnings for each education type. For 1979, I obtain $\hat{\sigma}^2_{\nu,H} = 0.0058$, $\hat{\sigma}^2_{\nu,L} = 0.0042$, $\hat{\sigma}^2_{\gamma,H} = 0.162$ and $\hat{\sigma}^2_{\gamma,L} = 0.1732$. I directly feed the estimated variances of the persistent shocks, $\hat{\sigma}^2_{\nu,H}$ and $\hat{\sigma}^2_{\nu,L}$, into the model and calibrate the variances of the temporary shock inside the model to match the Gini index for cross-sectional earnings equal to 0.30. When doing so, I impose an additional restriction that the ratio of the variances between the two types remains at $\sigma^2_{\gamma,H}/\sigma^2_{\gamma,L} = 0.935$, consistent with the empirical estimates of $\hat{\sigma}^2_{\gamma,z}$ obtained above. The calibrated variances of the temporary shock that follow from this procedure are the following: $\sigma^2_{\gamma,H} = 0.0056$ and $\sigma^2_{\gamma,L} = 0.006$.

I proceed in the exact same way for 2017 (matching the Gini index for pre-tax earnings equal to 0.39). I obtain $\sigma^2_{\nu,H} = 0.007$, $\sigma^2_{\nu,L} = 0.0065$, $\sigma^2_{\gamma,H} = 0.0126$, $\sigma^2_{\gamma,L} = 0.061$. These estimates imply that the variance of the persistent component has increased by 2.17 times for college graduates and 1.45 times for high school graduates between 1979 and 2017. The corresponding numbers for the variance of the temporary component are 1.25 and 1.08.[32]

Social Security

Throughout the calibration section, I take the replacement rate schedule in (13) as given and estimate its parameters $\boldsymbol{\alpha}$ outside the model. While Appendix C.1 provides the reader with all the details, I only report the obtained estimates: $\boldsymbol{\alpha}^{data}_{1979} = (0.45, 1.69)$ and $\boldsymbol{\alpha}^{data}_{1979} = (0.50, 1.48)$. Figure 2 contrasts the implied replacement rate schedules in 1979 and 2017. One can immediately detect a slight upward shift in the replacement rate of a household whose average lifetime earnings at retirement are precisely equal to the economy-wide average taxable earnings. This agent's normalized average lifetime earnings are given by $\hat{e} = 1$, marked by a vertical dashed line in the figure. Simultaneously, the replacement rates rise for all workers whose average lifetime earnings are below 35 percent of the economy-wide average taxable earnings, whereas the opposite occurs to the remaining retirees.

Regarding Social Security's remaining parameters, the mandatory retirement age, J^R, is 41 (real-life age 65) in the 1979 calibration. For 2017, the retirement age increases by one

[31] Also using the PSID, Guvenen (2009) estimates ρ_z separately for a sample of college graduates and a sample of high school graduates and finds that the estimates between the two samples differ only slightly. Particularly, he obtains $\rho_H = 0.979$ and $\rho_L = 0.972$.

[32] Using a pooled sample of households, Heathcote et al. (2010a) estimate that the variance of the persistent component has increased by 2.83 times between 1979 and 2000 (their most recent estimates), while the same number for the temporary component is 1.49. See their online appendix for more details.

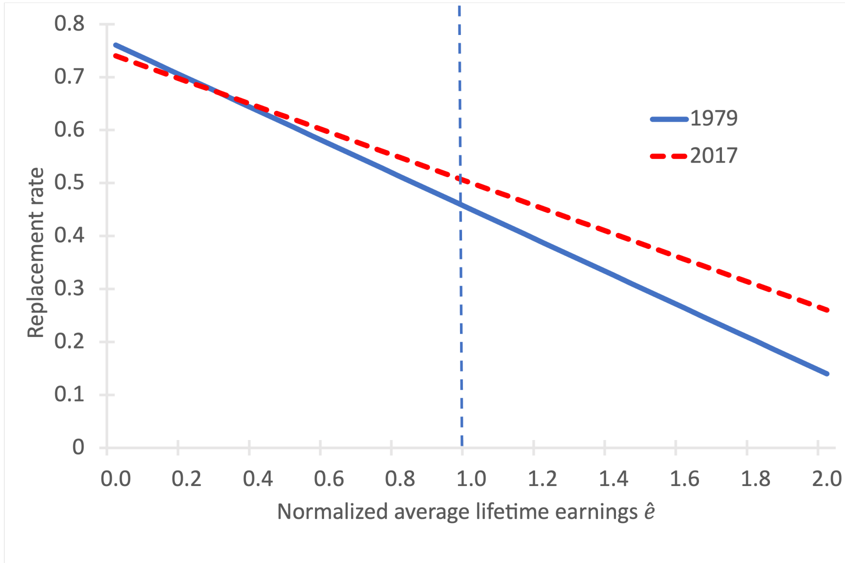

Figure 2: CALIBRATED REPLACEMENT RATE SCHEDULE.

Notes: The figure shows the calibrated replacement rate schedule in 1979 and 2017. The replacement rate schedule in the data is approximated using the specification in (13) and its parameters are estimated using a non-linear least squares estimator. Appendix C.1 elaborates on the estimation procedure.

year.[33] Furthermore, I calibrate cap_{SS} to match the ratio of the maximum taxable earnings threshold (adjusted for the number of earners in a household, see Appendix C.1) and the average taxable earnings in the data. Whereas this ratio was 1.79 in 1979, it increased to 2.32 in 2017. The equilibrium value of the Social Security tax rate, τ_{SS}, which was not a calibration target, is 8.99 percent in 1979 and 10.45 percent in 2017. Its empirical counterpart is the Old-Age and Survivors Insurance tax (sum of employers and employees' portions). With 8.66 percent in 1979 and 10.03 percent in 2017, it comes surprisingly close to the model-generated values.[34] Finally, I set the fraction of the Social Security deficit in GDP, d_{SS}, to 0.1 percent in 1979 and 0.47 percent in 2017, consistent with the data.[35]

[33]There are roughly 80 percent of retired household heads whose age is above or equal J^R in the 1980 and 2018 CPS extracts.

[34]See Table 4.B4 in Social Security Administration (2019) on the time series data for the Old-Age and Survivors Insurance tax.

[35]The data are publicly available at `https://www.ssa.gov/oact/STATS/table4a1.html`. As for the Social Security Trust Fund's expenditures, I take benefit payments and exclude the administrative expenses and transfers to the Railroad Retirement program. As for the Trust Fund's revenues, I take payroll tax contributions and exclude income from taxation of benefits, General Fund Reimbursements, and net interest. The data for the nominal GDP comes from FRED available at `https://fred.stlouisfed.org`.

Utility and technology

The capital share in production, θ, is chosen to match the average ratio of capital income in GDP, whereas the depreciation rate of capital, δ, is calibrated to match the average ratio of investment relative to GDP. I calibrate the discount factor, β, inside the model to match the capital-to-output ratio of 3.2 and 3.7 in the initial and final steady states, respectively. The respective targets were computed using National Income and Product Accounts.[36]

The instantaneous utility function $u(c, l)$ in (11) is a constant relative risk aversion function:

$$u(c, l) = \frac{[c^{\gamma}(1 - l)^{1-\gamma}]^{1-\sigma}}{1 - \sigma}, \tag{20}$$

where σ controls the degree of relative risk aversion and γ is the relative weight on consumption. I keep $\sigma = 2$ throughout all experiments. I calibrate γ inside the model, so that working-age agents spend, on average, 42 percent of their discretionary time endowment in 1979 and 43 percent in 2017 on work.[37]

The bequest motive is modeled as in De Nardi (2004) with the following functional form:

$$\phi(a) = \phi_1(1 + a/\phi_2)^{1-\sigma},$$

where ϕ_1 reflects the agent's concern about leaving bequests, whereas ϕ_2 measures the extent to which bequests are a luxury good. Following De Nardi (2004), I set $\phi_1 = -9.5$ and $\phi_2 = 11.6$ and keep it fixed across the steady states.

I set the degree of altruism to $\chi = 0.5$ and keep it fixed throughout all experiments. The choice of χ does affect the level of the estimated Pareto weights in each steady state. However, it does not affect the *relative* change in the Pareto weights across time.[38]

Aging

The key parameter that drives the longevity dynamics in the model is the conditional survival probability rate, $\psi_{z,j}$. Apart from rising earnings inequality, improved longevity

[36] I follow the procedure described in Hosseini and Shourideh (2019) (see their online Supplement, section S5, which is available at `https://www.econometricsociety.org/sites/default/files/ecta200042-sup-0001-Supplement.pdf`). For the initial steady state, I compute the averages of the relevant variables over the period 1970–1980 and for the final steady state I compute the averages over the period 2010–2018.

[37] According to the CPS data, household heads worked, on average, 2,161 hours in 1979. Given the total annual disposable time of 5,096 hours (see footnote 30), we have: 2,162/5,096 $\approx$ 42 percent. The corresponding target for 2017 is 2,075/5,096 $\approx$ 41 percent.

[38] See footnote 53.

is one of the most significant structural transformations that the United States economy experienced during the past four decades. According to Bell et al. (1992), the life expectancy of a Social Security covered worker increased at all ages during 1970–2010. Most significantly, longevity increased for a 25-year-old worker (7 years).[39] However, there is large heterogeneity in life expectancy. More specifically, the empirical literature documents a very robust finding that mortality rates tend to drop with income, education, and other socioeconomic status measures in the United States.[40] Moreover, the literature has argued that these differentials have been rising over time. Elo and Preston (1996), Meara et al. (2008), and Bound et al. (2014) find this pattern for educational differentials, while Waldron (2007) finds this pattern for earnings using Social Security Administrative data.[41]

In this paper, I incorporate the *ability*-specific differences in mortality.[42] However, I am unaware of any consistent data for the 1970s.[43] Hence, I assume that no such differences exist in the 1979 calibration. To parameterize the conditional survival rates, I directly feed into the model the data from Bell et al. (1992) for 1970. For the 2017 calibration, I rely on the data by Bound et al. (2014). The authors use the data from the National Vital Statistics System and the Census for 1990–2010 to compute life expectancies for a 25-year-old individual by education at five-year intervals (25, 30, etc). I focus on their most recent estimates calculated for 2010. Their results imply that a 25-year-old individual with a college degree or higher expects to live, on average, 54 years. In contrast, the same statistic for an individual with a lower education level is only 51 years. The way I utilize these data in the model is as follows. Appendix C.2 provides all the necessary details, so I summarize the procedure only briefly. I specify each ability group's survival rates as a Gompertz force of mortality function.

[39]See their Table 6. The authors report life expectancy separately for female and male Social Security covered workers. Since my model is agnostic about household's gender, I compute the weighted average between the reported life expectancy using the population share of female heads consistent with the CPS. I apply this procedure to all the estimates I report from Bell et al. (1992) in the paper.

[40]See Bound et al. (2014) for references.

[41]Meara et al. (2008) provided one potential explanation for the observed trends. They show empirical evidence that smoking rates and death rates caused by smoking-related diseases drop more significantly over time for higher socioeconomic groups than for the lower ones.

[42]In my model, income is an endogenous variable. Letting mortality depend on income is problematic because agents will find it optimal to affect their expected lifespan by adjusting their behavior. By contrast, the ability is a characteristic outside of the agent's control in the model.

[43]The only available empirical evidence I am aware of is Meara et al. (2008). The authors use the National Longitudinal Mortality Study (NLMS), which follows the March CPS households through subsequent deaths during 1981–85. According to their estimates, life expectancy at age 25 (averaged across gender) is 77 years for an individual with a high school degree or below and 79 years for an individual with at least some college education. Given that this evidence points to minor differences in mortality, I ignore them in the 1979 calibration.

I estimate the parameters of the Gompertz function by matching several empirical targets. First, I require that the model-based profile of survival rates averaged across the two ability types matches the life expectancy profile reported by Bell et al. (1992). Second, I require the type-specific age profiles of survival rates in the model to match the empirical counterparts reported by Bound et al. (2014).

Given the estimates of $\psi_{z,j}$, I calibrate the birth rate, n, outside the model to match the dependency ratio in each steady state consistent with the CPS data. I target the dependency ratio of 21 percent and 23 percent for 1979 and 2017, respectively. These numbers imply a reduction in the number of working-age households per retired household from 5 to 4.[44]

Income taxes

I take the specification of the income tax function from Heathcote et al. (2017):

$$\Lambda(e) = e - (1 - \tau_I)e^{1-\tilde{\tau}_I}, \tag{21}$$

where e is pre-government income and $\tau_I, \tilde{\tau}_I$ are parameters. $\tilde{\tau}_I$ determines the progressivity of the income tax schedule, whereas τ_I governs the overall level of income taxation.[45] To parameterize $(\tau_I, \tilde{\tau}_I)$, I rely on Wu (2020). The author uses the CPS data processed through the NBER's TAXSIM calculator to estimate the same income tax function as in my paper for two separate sample periods: 1978–1980 and 2014–2016. He finds that income tax progressivity, as measured by $\tilde{\tau}_I$, declines from 0.187 in the 1970s to 0.137 in the 2010s. I feed into the model the corresponding estimate of $\tilde{\tau}_I$ in each steady state. His estimated values of τ_I (0.155 in 1978–1980 and 0.078 in 2014–2016) correspond to the average tax rates faced by a median-income household. Hence, I calibrate τ_I inside the model to match the average

[44] To compute the dependency ratio in the CPS, I define a working-age household as a household whose head is of real-life age 25–64 in 1979 and 25–65 in 2017. A retired household is a household whose head is of age 65–85 in 1979 and 66–85 in 2017. These numbers are consistent with the choice of J^R and J.

[45] The income tax schedule is referred to as *progressive* (*regressive*) if the ratio of marginal to average tax rates is larger (smaller) than 1 for every level of income e. When $\tilde{\tau}_I > 0$, marginal income tax rates always exceed average rates, and the income tax system is progressive. Conversely, the tax system is regressive when $\tilde{\tau}_I < 0$. Given $\tilde{\tau}_I$, the second parameter, τ_I, shifts the tax function and determines the average level of taxation in the economy. The case with $\tilde{\tau}_I = 0$ implies that marginal and average tax rates are equal: the system is a flat tax with the tax rate $1 - \tau_I$.

Table 3: Targeted Moments.

Moment	**1979**	**2017**
Gini index earnings	0.30	0.39
Cap-to-earnings ratio	1.78	2.31
Average labor supply, %	42.41	40.72
Capital-to-output ratio	3.23	3.66
Dependency ratio, %	21.00	23.00
Income tax (median worker), %	15.05	7.80

Notes: The table displays the moments that the model matches exactly in each steady state (1979 and 2017). *Gini index for earnings* refers to the pre-tax earnings distribution. *Cap-to-earnings* is the ratio of the maximum taxable earnings threshold, cap_{SS}, to the economy-wide average taxable earnings, $\tilde{E}$. *Dependency ratio* shows the share of retired households to the working-age population. *Income tax* refers to the average tax rate faced by the household with median pre-government income.

income tax rate of a working-age agent with median labor income.[46]

Other taxes

I borrow the estimates for capital tax, τ_a, and consumption tax, τ_c, from Wu (2020) who applies the methodology developed by Mendoza et al. (1994) and Trabandt and Uhlig (2011). According to these estimates, τ_a falls from 38.4 percent in the late 1970s to 33.0 percent in the recent data. Congruently, τ_c declines from 5.3 percent to 4.1 percent.

The hospital insurance tax, τ_M, is 2.1 percent in 1979 and 2.9 percent in 2017, which is the sum of the employers' and employees' portions of the tax.[47] Consistent with the data, I set the maximum taxable earnings threshold for Medicare, cap_M, equal to cap_{SS} in the initial steady state. There has been no limitation on taxable earnings in the Medicare program starting from 1994, so I set cap_M to infinity, virtually eliminating the cap.

3.3. Model fit

As a first passthrough for how the model matches the data, Table 3 summarizes all the moments that I directly targeted during the calibration exercise. These moments were matched exactly.

[46]His definition of a worker's pre-government income comprises the wage and self-employment income; the capital income tax is estimated separately. This approach is consistent with my theoretical model's assumptions because only the labor income is subject to income taxation. Furthermore, Wu (2020) estimates the income tax function parameters using a sample of household-level data with household heads aged 23–62. In my model, the observation unit is also a household, though my sample selection criterion is less restrictive, as I keep all household heads of age 25–64.

[47]Table 4.B4 in Social Security Administration (2019).

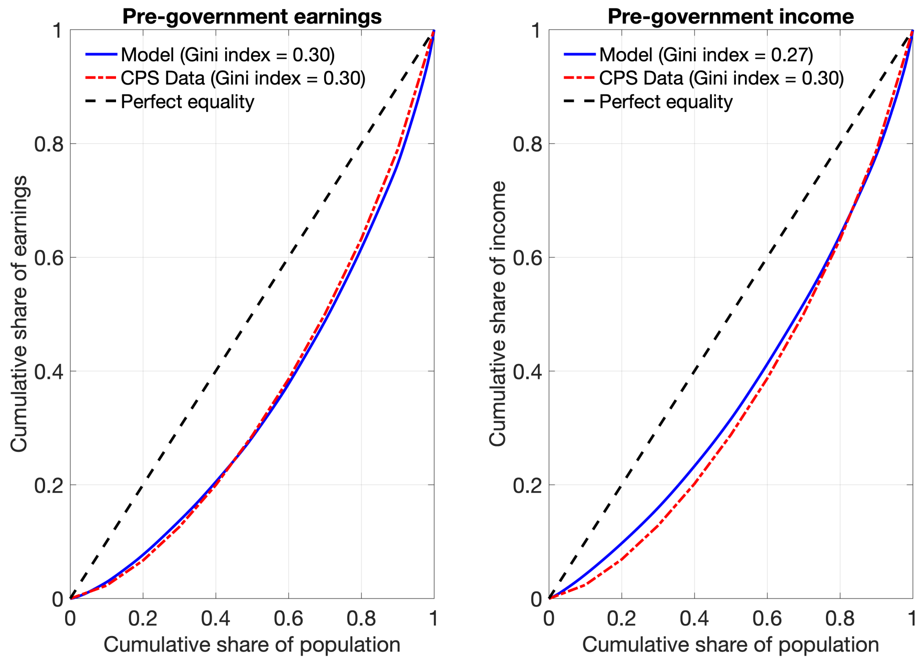

Figure 3: Fit of Earnings and Income Inequality in 1979.

Notes: The figure compares the fit of the Lorenz curves for the earnings distribution (left panel) and the income distribution (right panel) in 1979. The empirical Lorenz curves are computed using the household-level data from the 1980 CPS extract. The model-implied Lorenz curves are computed based on the steady state distribution under the 1979 calibration. Earnings and income are measured before taxes and government transfers.

By construction, the model matches the Gini index for pre-government earnings in both steady states. The earnings Gini rises from 0.30 in 1979 to 0.39 in 2017. To assess the model's ability to capture inequality in the data beyond the Gini index, I compute the model-based Lorenz curves for pre-government earnings and pre-government income and compare them with the CPS data.

The results are shown in Figure 3 for 1979 and Figure 4 for 2017. The left panel in both figures describes the earnings concentration, whereas the right panel displays the income concentration. One can see that the model fits the earnings inequality in both years very accurately. However, the model slightly overestimates the economic well-being of those households located in the lower tail of the income distribution. Nevertheless, the model's overall performance matching income inequality is surprisingly good, given that I did not target the Lorenz curves.

The final assessment of the model focuses on the distribution of college graduates by earnings. The joint distribution of education and earnings is essential because it controls the

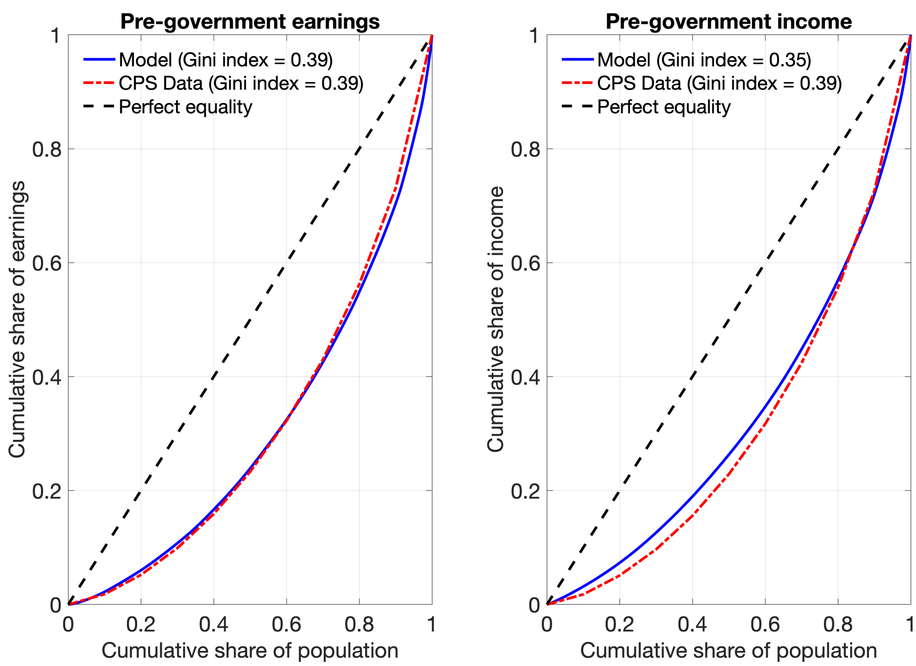

Figure 4: Fit of Earnings and Income Inequality in 2017.

Notes: The figure compares the fit of the Lorenz curves for the earnings distribution (left panel) and the income distribution (right panel) in 2017. The empirical Lorenz curves are computed using the household-level data from the 2018 CPS extract. The model-implied Lorenz curves are computed based on the steady state distributions under the 2017 calibration. Earnings and income are measured before taxes and government transfers.

Table 4: SHARE OF COLLEGE GRADUATES PER EARNINGS QUARTILE IN THE MODEL AND DATA.

Quartile	1979		2017	
	Model	Data	Model	Data
Q1	14.2	12.9	20.5	21.9
Q2	21.6	19.7	40.8	34.7
Q3	25.4	25.0	34.5	46.1
Q4	38.8	40.8	77.2	68.2

Notes: The table compares the shares of college graduates by quartiles of the earnings distribution in the data with the corresponding shares of high-ability agents in the model. The empirical moments are computed using the household level data from the 1980 and 2018 CPS extracts. The moments in the model are computed using the steady state distributions under the 1979 and 2017 calibrations.

forces that drive inequality (luck versus permanent ability differences). Table 4 conducts the assessment. Given that I did not directly target any of the moments presented in the table, the model achieves a fairly good fit.

4. Findings

The paper's findings are presented in the following order. Section 4.1 uncovers substantial heterogeneity in agents' preferences over the replacement rate schedule and builds intuition behind the results obtained in the remaining part of the paper. Section 4.2 addresses the paper's first question of how the government's preferences over redistribution have changed during the past four decades. Section 4.3 quantifies the welfare cost of the shift in government preferences by answering the paper's second question: How should Social Security have responded to earnings inequality, had the government's preferences remained constant?

4.1. Heterogeneity in preferences over replacement rate schedule

Consider the government's maximization problem in (16) and its first-order optimality conditions with respect to α_i with $i \in \{1, 2\}$:

$$\int_{\mathbf{x}} \omega(j, z; \boldsymbol{\kappa}_t) \frac{\partial V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}, F_t)}{\partial \alpha_i} dF_t(\mathbf{x}) = 0. \tag{22}$$

The equation above shows that the government's optimal choice in a given time t depends on three objects: the Pareto weights governed by the parameter vector $\boldsymbol{\kappa}_t$, the distribution of households over states $\mathbf{x}$, $f_t(\mathbf{x})$, and the first-order derivative of the household's value function with respect to policy α_i, $\partial V_t / \partial \alpha_i$. The last object is the focus of the current section.

Recall that the paper's first goal is to rationalize the *actual* replacement rate schedule. Therefore, I am interested in the derivative of the agents' value function evaluated at the *existing* policy, $\boldsymbol{\alpha}_t^{data}$. The sign and the magnitude of this derivative have an important economic interpretation. Consider first the derivative's sign. In all my simulations, the agents' value function is strictly concave in each of the two policy variables.[48] This means that if the derivative of the agent's value function with respect to α_i turns out to be negative, then the agents' most preferred policy α_i must lie below the observed policy. The opposite is true if the derivative is positive. Hence, the sign of the derivative allows me to detect agent groups that have major disagreements over α_i. Consider next the derivative's magnitude. It shows the size of the welfare gain (if the derivative is positive) or the size of the welfare loss (if the derivative is negative) to an agent induced by slightly increasing policy α_i. All else equal, the higher the derivative's magnitude for a given agent, the closer the government will move the policy toward this agent's most preferred policy.

Since the model exhibits multiple dimensions of heterogeneity, it is not feasible (but also not necessary) to explore the variations in welfare along each of these dimensions. Instead, I focus on two dimensions: age and ability. I argue that they are strong determinants of agents' preferences over the replacement rate schedule. Not coincidently, the Pareto weights in the government optimization problem are specified as a function of the same agent's characteristics.

My analysis proceeds as follows. After computing $\partial V_t / \partial \alpha_i$ for each individual state, $\boldsymbol{x}$, I calculate the average value within pre-specified subgroups. Along the age dimension, I build three groups of agents: *young* (real-life age 25–39), *middle-aged* (40–64), and *retired* (64 and older).[49] Concerning ability, there are only two groups. Table 5 displays the average derivative for each subgroup for the 1979 calibration.[50] Qualitatively, all the arguments presented in this section hold also for the 2017 calibration.

4.1.1. Marginal welfare effects of α_1

Consider first the derivative of the agent's welfare function with respect to α_1, evaluated at the true policy $\boldsymbol{\alpha}_{1979}^{data}$ (middle column of Table 5). Observe that there is no disagreement between high-ability and low-ability agents inside each age group regarding which direction

[48] I do not have a formal proof of the strict concavity of the agent's value function with respect to policies α_1 and α_2; however, in the calibrated economy that I solve numerically, I check that the indirect utility function satisfies this property for every state $\boldsymbol{x}$.

[49] This partitioning choice is not crucial for the results since it only serves the illustrative purpose.

[50] Appendix B.3 describes the numerical procedure to compute $\partial V_t / \partial \alpha_i$.

Table 5: Marginal Welfare Effects.

	Policy α_1	**Policy** α_2
Young:		
– High-ability	-4.35	-0.42
– Low-ability	-2.66	0.50
Middle-aged:		
– High-ability	2.35	-0.55
– Low-ability	3.88	0.38
Retired:		
– High-ability	3.39	-0.30
– Low-ability	4.31	0.23

Notes: The table shows the partial derivatives of the value function V_{1979} with respect to the average replacement rate, α_1, and progressivity, α_2. V_{1979} is obtained from solving the model for the steady state equilibrium associated with $\boldsymbol{\alpha}_{1979}^{data}$ under the 1979 calibration. After obtaining the individual derivatives for each state $\mathbf{x}$, I calculate the average derivative for subgroups defined by age and ability. Along the age dimension, I build three groups: young (real-life age 25–39), middle-aged (40–64), and retired (above 64). I use the initial steady state distribution, f_{1979}, to average the individual derivatives within each subgroup.

to shift α_1 because the derivatives' signs are identical for both ability types. Next, the value function for the young as a group must be downward sloping because the derivative is negative for each ability type. This observation indicates that the young agents' most preferred average replacement rate must lie below the data's actual level. The opposite is true for middle-aged and retired agents, for whom the derivative has a positive slope. Thus, there is a disagreement between the young, on the one hand, and middle-aged and retired agents, on the other hand, regarding the average replacement rate level. This inter-generational conflict over α_1 will allow me to identify the age bias parameter $\kappa_{1,1979}$ in the next section.

What economic forces drive this inter-generational conflict? To answer this question, I inspect how the key determinants of the agent's welfare depend on policy α_1. I start with the model economy calibrated for 1979 and analyze how the equilibrium interest rate, wage rate, and the Social Security tax change when I vary α_1. Throughout the experiment, I fix the pension system progressivity at $\alpha_{2,1979}^{data}$. Besides prices and the Social Security tax, I also analyze the redistributive effect of pension benefits. More specifically, I ask: By how much percent do pension benefits reduce the market income inequality (measured by the Gini index) among retired agents?

Figure 5 shows the results. The values of all variables, except the pre-tax wage, are shown as percentage point deviations from the initial steady state in 1979; deviations in the pre-tax

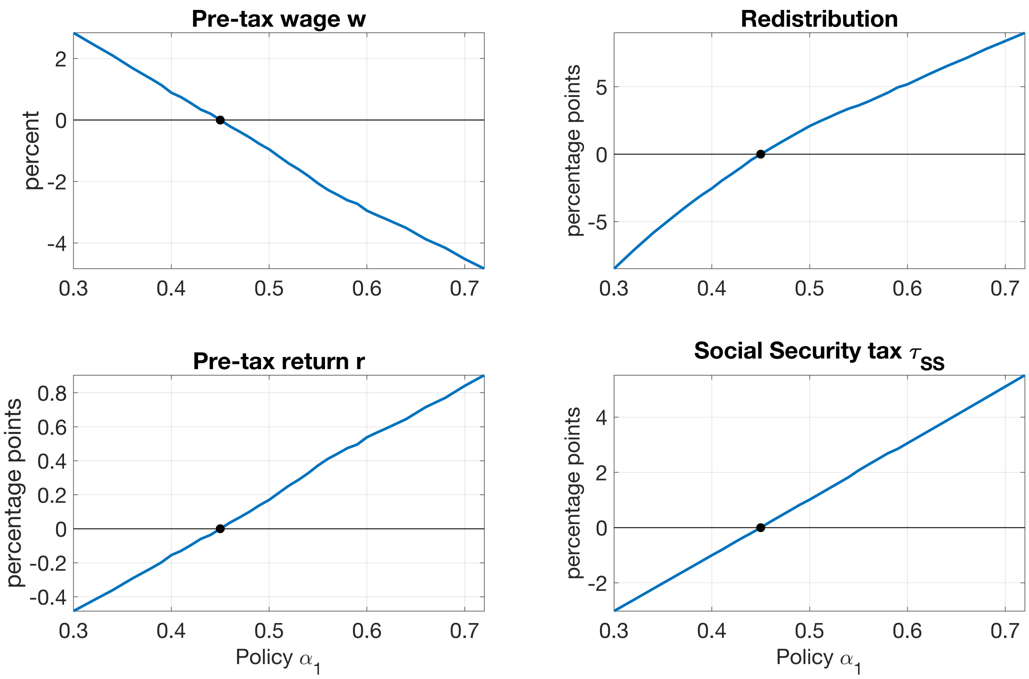

Figure 5: Model Variables as a Function of α_1.

Notes: The figure shows the relationship between the key model variables and the average replacement rate α_1. To construct the figure, I solve the model for a set of steady state equilibria under the 1979 calibration of all parameters, except for α_1. In all computations, pension system progressivity is fixed at $\alpha_{2,1979}^{data}$. On the vertical axis, I plot the equilibrium values of the pre-tax wage w, the pre-tax return on saving r, the Social Security tax rate, τ_{SS}, and the redistributive effect of pensions. The latter statistic is defined as the percentage reduction in market income inequality, measured by the Gini index, due to pension benefits. Values for w are shown in percent deviations, whereas the values of all other variables are shown in percentage point deviations from the baseline model in 1979. By construction, all variables are equal to zero at $\alpha_{1,1979}^{data}$.

wage are measured in percent. By construction, all depicted variables are equal to zero at $\alpha_{1,1979}^{data}$ (marked by a dot in the figure). Note that there is an entire transitional path of all endogenous variables from the initial steady state associated with $\alpha_{1,1979}^{data}$ to a new steady state under a given policy choice α_1. The figure shows the values of the variables in the *final steady state*.

Consider an increase in α_1. As the entire replacement rate schedule shifts upward, each retiree in the economy enjoys a larger pension annuity. As total pension entitlements rise, the Social Security tax rate (bottom-right panel) has to adjust upward to satisfy the government budget constraint. Since the public pension system crowds out private savings, the aggregate capital stock and capital intensity (not shown in the figure) decline. This leads to a reduction in the pre-tax wage w (top-left panel) and an increase in the pre-tax interest rate r (bottom-left panel). The after-tax wage, $(1-\tau_{SS})w$, declines, as well.

As previously pointed out, the young would like to reduce α_1 *below* the actual level. Two motives explain this fact. First, a smaller size of the pension system would result in higher pre-tax and after-tax wages through the general equilibrium. The young like the prospect of having higher after-tax labor incomes throughout their working careers since these agents rely substantially on earnings as their main source of income. Second, the economy is dynamically efficient in all my simulations, meaning that the young would receive a higher after-tax return on their savings than the Social Security's implicit return. Note, however, that the young agents' most preferred policy α_1 is strictly above zero. Workers face uncertainty in their future earnings, so they value the insurance against low realizations of earnings shocks that Social Security provides.

Contrary to the young, retired agents would benefit from *raising* the average replacement rate above the actual level because this would immediately increase retirees' pension benefits. Furthermore, the after-tax return on saving would become higher through the general equilibrium, making retired agents even happier, given that savings are their only source of market income. Middle-aged agents have to trade off falling after-tax wages against higher future replacement rates. Given my partitioning choice for the middle-aged group, which includes agents aged 40–64, the marginal welfare effect from increasing the policy turns out to be positive for this group, indicating that the benefit side dominates the cost.

As I have shown above, the Social Security tax rate is sensitive to the changes in the average replacement rate. For this reason, the policy variable α_1 controls the *inter-generational* redistribution in the model, i.e., the redistribution between workers and retired agents. As I will illustrate immediately, pension system progressivity has an insignificant impact on the Social Security tax rate and prices in the calibrated model economy.

4.1.2. Marginal welfare effects of α_2

Next, I discuss the marginal welfare impact of the pension system progressivity. The last column of Table 5 summarizes the marginal welfare effects of policy α_2. It is immediately noticeable that there is a disagreement between low-ability and high-ability agents within each group regarding which direction to shift policy α_2. Particularly, low-ability agents prefer, on average, to increase progressivity above the actual level, whereas the opposite is true for high-ability agents. The distributional conflict between high-ability and low-ability agents over policy α_2 will allow me to identify the ability bias parameter $\kappa_{2,1979}$.

What economic forces determine the agents' most preferred choice? Before answering this question, it is instructive to consult Figure 6, which plots the steady state values for the key model variables as a function of pension system progressivity. As one can see from the figure, variable α_2 has a relatively small quantitative impact on the Social Security tax rate and prices compared to policy α_1.[51] Instead, it controls the distribution of pension benefits within the pool of retired agents while roughly maintaining the overall fixed size of the pension system. Hence, α_2 governs the degree of *intra-generational* income redistribution in the model, i.e., the redistribution from retired agents with high average lifetime-earnings to the retirees with low average lifetime earnings.

Agent's preferences over intra-generational redistribution depend on the position of this agent's average lifetime earnings, $\bar{e}$, relative to the economy-wide average taxable earnings, $\tilde{E}_t$, in the last period before the agent's retirement. Since there is no aggregate risk in the model, all agents know $\tilde{E}_t$ with certainty at any point in time. Moreover, $\bar{e}$ remains constant during retirement, as can be seen from its law-of-motion in (8). Thus, the retired agent's relative position in the distribution of lifetime earnings, $\hat{e} = \bar{e}/\tilde{E}_t$, remains constant and known over time. As we can see from the replacement rate schedule in (13), an increase in α_2 reduces pensions of all agents with $\hat{e} > 1$. Thus, all retirees with $\hat{e} > 1$ must prefer $\alpha_2 < 1$. The opposite is true for retirees with $\hat{e} < 1$. Contrary to retirees, working-age agents face idiosyncratic labor productivity risk. Hence, their most preferred progressivity depends on the *expected* relative position in the distribution of lifetime earnings when these agents enter retirement.

What do these observations imply for the agents' most preferred policy by their age and ability? In the calibrated model economy, almost all low-ability retired agents (90.3 percent) turn out to have average lifetime earnings below the economy-wide average earnings. The ability composition changes substantially in the pool of retirees with average lifetime earnings above the economy-wide average earnings, where the share of low-ability agents drops to 58.6 percent. Hence, on average, low-ability retirees prefer a more progressive pension system (i.e., a higher α_2) than high-ability retirees because most of these agents are poor in terms of their lifetime earnings. Since high-ability workers face a deterministic wage premium during their entire working careers, the agent's ability becomes a dominant predictor of her expected lifetime earnings at retirement. Thus, high-ability agents can expect to enter retirement with average lifetime earnings above the economy-wide average with higher probability. Ex-post, i.e., after the realization of idiosyncratic risk, this is indeed what happens in the model. This intuition explains the nature of the intra-generational conflict between low-ability and

[51] The elasticity of the Social Security tax rate with respect to policy α_1 is 1.08, whereas it is -0.03 for α_2.

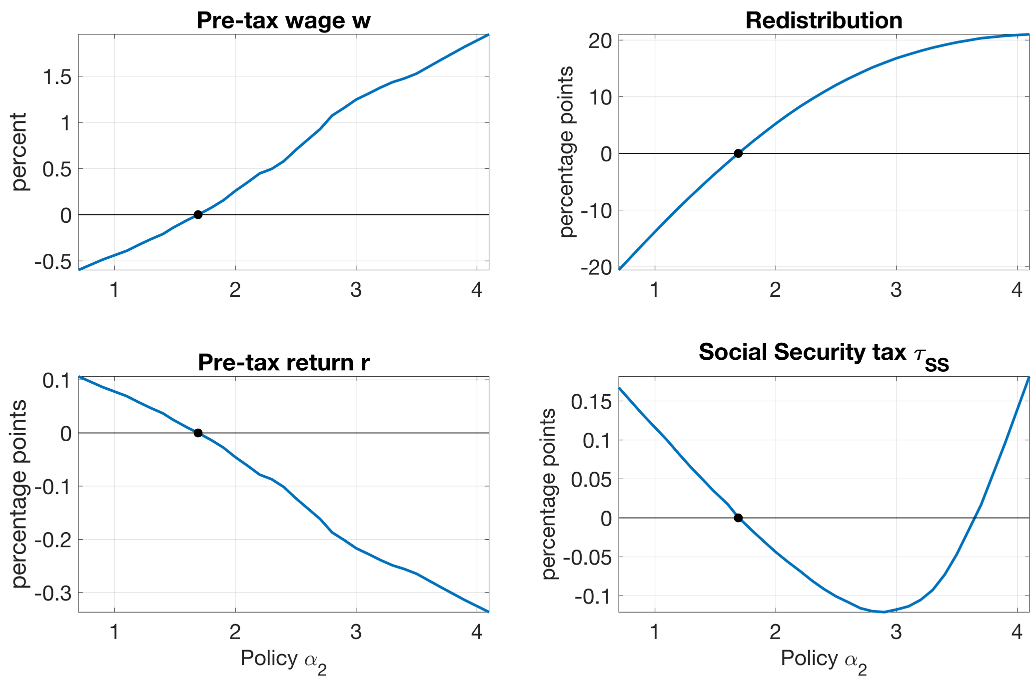

Figure 6: MODEL VARIABLES AS A FUNCTION OF α_2.

Notes: The figure shows the relationship between the key model variables and the statutory progressivity α_2. To construct the figure, I solve the model for a set of steady state equilibria under the 1979 calibration of all parameters, except for α_2. In all computations, the average replacement rate is fixed at $\alpha_{1,1979}^{\mathtt{data}}$. On the vertical axis, I plot the equilibrium values of the pre-tax wage w, the pre-tax return on saving r, the Social Security tax rate, τ_{SS}, and the redistributive effect of pensions. The latter statistic is defined as the percentage reduction in market income inequality, measured by the Gini index, due to pension benefits. Values for w are shown in percent deviations, whereas the values of all other variables are shown in percentage point deviations from the baseline model in 1979. By construction, all variables are equal to zero at $\alpha_{2,1979}^{\mathtt{data}}$.

high-ability agents highlighted in Table 5.

4.2. How have the government's preferences changed?

Based on my calibration (Table 2), the actual average replacement rate increased by 11 percent, whereas actual progressivity declined by 12 percent between 1979 and 2017. In the model, two distinct forces are capable of rationalizing this development. Conditional on the calibrated model parameters for 2017, the government might find it optimal to adjust the replacement rate schedule because its tastes over redistribution have changed over time. This is the first force which I will discuss in the current section. Simultaneously, the government might find it optimal to adjust the replacement rate schedule in 2017 because it faces a new set of model parameters, conditional on its preferences over redistribution from 1979. This is the second force that I will turn my attention to in the next section.

Recall that in the model, government tastes for redistribution are captured by the Pareto weights $\boldsymbol{\omega}_t$ (eq. 17) specified as a function of two parameters: the age bias ($\kappa_{1,t}$) and the ability bias ($\kappa_{2,t}$). In the previous section, I showed that there is a major disagreement *between* generations over the average replacement rate in the calibrated model economy. At the same time, there is large heterogeneity in welfare between high-ability and low-ability agents *within* each age group over the degree of pension system progressivity. In the current section, I will exploit these two distributional conflicts to identify the Pareto weight parameters $\boldsymbol{\kappa}_t$ consistent with the actual Social Security policy in 1979 and 2017.

Table 6 summarizes this section's main findings. Columns *Baseline* document the calibrated parameters of the replacement rate schedule under each calibration. The last two rows of the table show the estimated values of $\boldsymbol{\kappa}_t$ consistent with the government optimally choosing the calibrated replacement rate schedule.[52]

For the 1979 calibration, I obtain a negative value for $\kappa_{1,1979}$ and the estimate of $\kappa_{2,1979}$ above one. These results have the following implication. For the model to rationalize $\boldsymbol{\alpha}_{1979}^{\mathtt{data}}$, the government must have put a larger weight on younger agents, conditional on their ability. Simultaneously, the government must have attached a larger weight on high-ability agents compared with low-ability agents, conditional on their age. Under the 2017 calibration, the age bias switches its sign and becomes positive, which suggests that the government in 2017 must have put a higher weight on older agents, conditional on their ability. At the same time, the ability bias continues to be positive and increases in magnitude, which means that the government must have assigned an even higher relative Pareto weight to high-ability agents,

[52] Appendix B.3 explains in detail the numerical algorithm.

Table 6: Actual and Utilitarian Policies in 1979 and 2017.

	1979		2017
	Equal weights	Baseline	Baseline
Average rep. rate $\alpha^{\star}_{1,t}$	0.58	0.45	0.50
Progressivity $\alpha^{\star}_{2,t}$	1.90	1.69	1.48
Soc. Sec. tax $\tau_{SS,t}$, %	11.6	9.0	10.5
Redistribution, %	62.1	53.0	31.0
Age bias $\kappa_{1,t}$	0	−0.86	0.64
Ability bias $\kappa_{2,t}$	1	2.77	5.73

Notes: The table shows a set of variables computed in different experiments. *Equal weights* refer to the utilitarian policy in 1979. *Baseline* refer to the benchmark model economy under the 1979 or 2017 calibration. For each experiment, the table records the optimal replacement rate schedule, $\boldsymbol{\alpha}^{\star}_t$, the Social Security tax rate as well as the measure of income redistribution, and the Pareto weight function parameters, $\boldsymbol{\kappa}_t$, that give rise to $\boldsymbol{\alpha}^{\star}_t$. The Social Security tax and the redistribution statistic are computed based on the final steady state associated with a given optimal policy $\boldsymbol{\alpha}^{\star}_t$. The redistribution measure shows the percent that pension benefits reduce market income inequality (measured by the Gini index) among retired agents.

conditional on their age, compared with the one in 1979.

It is helpful to put the estimates of $\boldsymbol{\kappa}_t$ into perspective. Figure 7 visualizes the shift in the age bias estimate $\kappa_{1,t}$. More specifically, it contrasts the implied age-profile of Pareto weights in 1979 and 2017, keeping the agent's ability fixed. To ease the comparison across periods, I normalize the Pareto weights by the weight attached to a 65-year-old agent in each period. As one could expect from the sign of the obtained estimates, the Pareto weights decrease in the agent's age in 1979, while the opposite happens in 2017. Next, consider a 25-year-old household. In 1979, the government was willing to trade off one util of this agent against more than two utils of a 65-year-old household with the same ability. This tradeoff is virtually reversed in 2017. Overall, the estimates of $\kappa_{1,t}$ suggest that the government has become more willing to tolerate redistribution from young workers toward retired households.

Since the ability bias, $\kappa_{2,t}$, enters multiplicatively into the Pareto weight function in (17), its estimate directly captures the Pareto weight that the government assigns to a high-ability agent relative to a low-ability agent of the same age. Thus, $\kappa_{2,1979} = 2.77$ means that the Pareto weight on a high-ability household is roughly three times larger than the weight on a low-ability household of the same age. Since $\kappa_{2,t}$ increases from 2.77 to 5.73, the Pareto

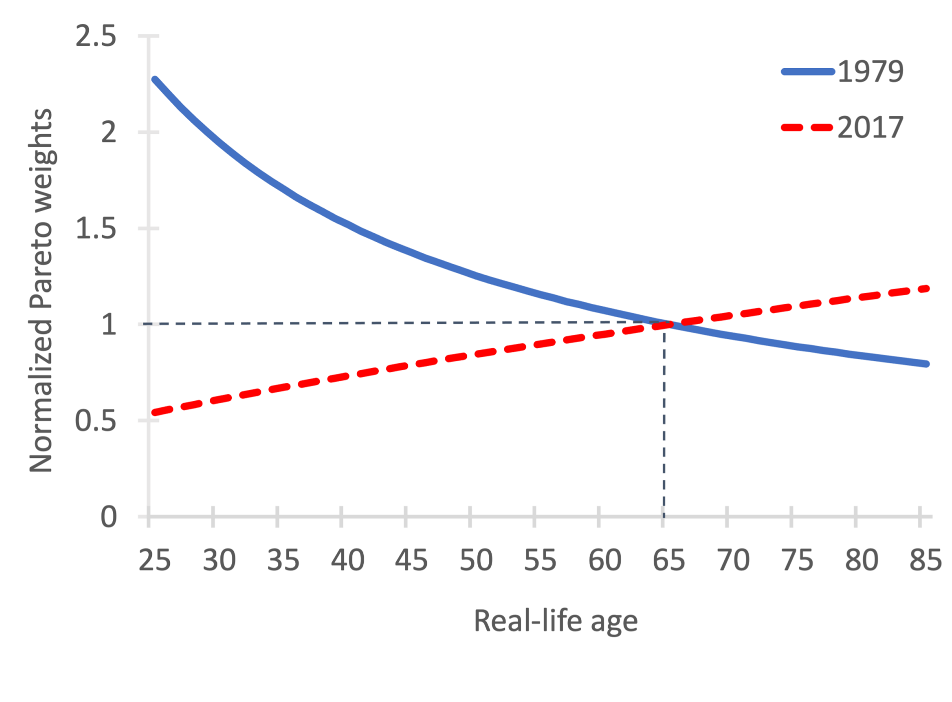

Figure 7: Pareto Weights By Age.

Notes: The figure shows the age-profile of Pareto weights, $j^{\kappa_1,t}$, conditional on the agent's ability, in 1979 and 2017. The weights are normalized by the Pareto weight attached to a 65-year-old agent in each period.

weight on a high-ability household almost doubles over time.[53] According to this finding, the government has become less willing to redistribute from high-ability to low-ability households.

Admittedly, the model is salient about the forces underlying the identified changes in the government's preferences. In Appendix D, I present empirical evidence regarding the changes in the voter turnout rates by education and age in Congressional elections in the United States. This evidence is largely in congruence with the model predictions.

The remaining part of this section provides some intuition behind the results obtained above. Consider the 1979 calibration of the model. Qualitatively, all the arguments to follow also apply for the Pareto weights identification in the 2017 calibration of the model. To establish a useful benchmark for comparison, I compute and report in column *Equal weights* in Table 6 the optimal *utilitarian* replacement rate schedule in 1979.[54]

As we can see from the table, the utilitarian government chooses the average replacement rate equal to 0.58 and the degree of pension system progressivity equal to 1.9. Hence, the utilitarian solution implies a 29 percent higher level of replacement rates and a 12 percent higher progressivity compared to the actual policy. The discrepancy between the two solutions can be summarized using the consumption equivalent variation (CEV) measure. It shows how much (in percent) the agents' consumption has to be increased in all future periods and contingencies (keeping their leisure unchanged) in the baseline model under the 1979 calibration so that their expected utility equals that under the utilitarian policy. Consistent with the government optimization problem, I compute this measure among all living agents in 1979. The utilitarian policy leads to a welfare gain equivalent to a 0.98 percent consumption increase.

To further quantify the discrepancy between the two policies, I also report the Social Security tax rate and the redistributive effect of pensions associated with each policy, where the latter statistic measures the percent reduction in market income inequality (measured by the Gini index) due to pension benefits. Similar to the remark made in the previous section, there is an entire transitional *path* of all endogenous variables from the initial steady state with policy $\boldsymbol{\alpha}_t^{data}$ to the final steady state associated with the optimal utilitarian policy. Therefore, in the table, I report the values of the Social Security tax rate and the redistribution measure obtained in the *final* steady state.

[53] As already pointed out in the calibration section, the estimates of Pareto weights depend on the degree of altruism, χ. I conducted a robustness exercise setting $\chi = 0$. The obtained estimates of the Pareto weight function are: $\boldsymbol{\kappa}_{1979} = (-3.70, 1.60)$ and $\boldsymbol{\kappa}_{2017} = (-2.14, 3.53)$. Thus, χ does not affect, qualitatively, the detected shift in the Pareto weights toward older and high-ability households.

[54] Appendix B.4 describes in detail how the utilitarian solution is obtained.

One can see that the steady state value of the Social Security tax rate implied by the utilitarian solution is 2.6 percentage points higher compared with the data. This difference is explained by the fact that the utilitarian government sets a higher average replacement rate which requires additional financing due to the government budget constraint (15). Concurrently, income redistribution among retired agents rises by 9 percentage points compared with the data because the utilitarian government chooses a more progressive pension system.

By construction, any discrepancy between the utilitarian and the actual policies is absorbed by the Pareto weights ω_t. In the previous section, I demonstrated that the young are the only group in favor of reducing α_1 below the actual level in the data; hence, the Pareto weight attached to this group must be higher than in the utilitarian case for the model to match $\alpha_{1,1979}^{data}$. Furthermore, the utilitarian solution leads to the degree of progressivity above the empirically observed level in 1979. Compared with the utilitarian case, a larger Pareto weight on high-ability agents is required because, as we saw in the previous section, high-ability agents prefer a less progressive pension system than in the data.

4.3. How should have Social Security responded to inequality?

As opposed to the previous discussion, this section assumes that the government's preferences over insurance and redistribution have not changed over time. I maintain this assumption throughout this section to disentangle the government's optimal response to parameter changes from any shifts in the government's preferences. This allows me to quantify the welfare cost of the shift in government preferences.

The economy is assumed to be in a steady state equilibrium at time $t = 2017$. The initial Social Security policy is $\boldsymbol{\alpha}_{1979}^{data}$ and the initial distribution of agents is given by the steady state distribution that arises in the steady state equilibrium under policy $\boldsymbol{\alpha}_{1979}^{data}$ and the following calibration of parameters. As already mentioned, I keep the Pareto weights fixed at their 1979 levels.[55] I compute the optimal replacement rate policy under two alternative calibrations. In the first case, I update all the model parameters to their calibrated values for 2017. In the second case, I update only those model parameters which drive earnings inequality, while leaving all the remaining parameters constant at their calibrated levels for 1979. This experiment allows me to quantify the relative contribution of earnings inequality to the optimal policy.

Table 7 displays the results. The values for $\boldsymbol{\alpha}^{\star}$ are shown in percent deviations relative to the baseline model in 1979, whereas the long-term values of the Social Security tax rate and

[55]The values of $\boldsymbol{\kappa}_t$ are taken from the column *Baseline* for 1979 in Table 6.

Table 7: Optimal Policy in 2017 in Counterfactual Experiments.

	All Changes	**Inequality**	**Premium**	**Risk**	**Graduates**
Average rep. rate $\alpha_1^\star$, %	−28.9	−11.1	−11.1	+2.2	−8.9
Progressivity $\alpha_2^\star$, %	+184.0	+59.8	+18.3	+24.3	−5.3
Soc. Sec. tax τ_{SS}, pt. pt.	−1.80	−1.84	−1.32	−0.09	−0.81
Redistribution, pt. pt.	−0.20	+4.86	+1.45	+4.37	−6.47

Notes: The table shows the optimal policy in 2017 in counterfactual experiments. All policies are computed under the identified Pareto weights from 1979 (Table 6, third column). The values for $\alpha_1^\star$ and $\alpha_2^\star$ are shown in percent deviations, whereas the values for the Social Security tax rate and redistribution are displayed in percentage point deviations from their respective values under the baseline calibration for 1979. All policies are computed under the Pareto weight parameters $\boldsymbol{\kappa}_{1979}$. The reference values for 1979 are displayed in the second column of Table 6. Redistribution measures the percent reduction in the Gini index for market income inequality due to pension benefits. τ_{SS} and the redistribution measure are both computed in the final steady state equilibrium associated with a given optimal policy. In each counterfactual, all parameters are fixed at their calibrated values for 1979, except for the following parameters that take on values from 2017: the age-efficiency profiles, $\{\zeta_{z,j}\}_{z,j=1}^{J^R-1}$ (*Premium*), the variances of the persistent and temporary shocks, $\{\sigma^2_{\gamma,z}, \sigma^2_{\nu,z}\}_z$ (*Risk*), and the share of high ability agents, λ_H (*Graduates*). *Inequality* combines all the changes from the previous three counterfactuals. *All Changes* update all the model parameters. See Table 2 for parameters updated.

the redistribution measure associated with the optimal policy are displayed in percentage point deviations relative to the same benchmark.

Before evaluating the marginal impact of earnings inequality on the optimal policy, I investigate Social Security's optimal response to a combined change in all model parameters (column *All Changes*). Under this scenario, the optimal average replacement rate declines by 28.9 percent. This measure allows the government to reduce the Social Security tax rate by 1.8 percentage points in the long run. Concurrently, the optimal progressivity skyrockets by 2.84 times. Despite this substantial increase, however, the government can only sustain roughly the same degree of redistribution as in 1979.

Next, I evaluate the welfare gain associated with this optimal policy. More specifically, I ask the following question. By how much percent does consumption of the agents in the baseline economy under the 2017 calibration have to increase in all future periods and contingencies (keeping their leisure unchanged) so that their expected utility equals that under the optimal policy reform? Consistent with the government optimization problem, I consider only those agents who are alive in 2017. The resulting welfare effect is 1.16 percent. This is the welfare cost that the agents in the baseline model would be willing to pay to move to an economy in which the government's preferences over insurance and redistribution have remained unchanged.

In the next step, I re-compute the optimal policy after updating only those model parameters responsible for the increased dispersion in earnings (column *Inequality* in Table 7).[56] According to the table, the government optimally chooses to reduce the average replacement rate by 11.1 percent and increase progressivity by 59.8 percent compared to 1979. Hence, qualitatively, the marginal effect of earnings inequality on the optimal policy in 2017 is the same as the combined effect of all model parameters that I described above. Quantitatively, the increased earnings dispersion explains 38.5 percent of the total decline in the average replacement rate and 32.5 percent of the total increase in progressivity.[57]

Below I will zoom onto the marginal impact of earnings inequality. To uncover the economic mechanisms that drive the results, I conduct a set of counterfactual experiments. Each experiment's set-up is identical to the *Inequality* exercise explained above with one important modification. Instead of updating all the parameters that drive inequality, I split them into three subgroups and study the marginal contribution of each subgroup to the optimal policy. In the first experiment, I update the deterministic age-profile of return to experience governed by $\zeta_{z,j}$. For the sake of brevity, I will refer to this case as *Premium* in the tables. The second experiment (*Risk*) accounts for the changes in the persistent and temporary shocks' variances, $\{\sigma^2_{\gamma,z}, \sigma^2_{\nu,z}\}_z$, only.[58] In the final experiment, I adjust the share of high ability agents, λ_H, only.

College premium

Consider the first counterfactual, in which I only update the profile of age-efficiency units, $\zeta_{z,j}$. Recall that this variable controls the deterministic portion of the agent's earnings. Conditional on all other characteristics, high-ability workers receive a wage premium over low-ability workers given by $\zeta_{H,j}/\zeta_{L,j}$. In the calibration section, Figure 1 illustrated that the wage premium shifts upward at each age between 1979 and 2017.

This development is responsible for a substantial rise in the dispersion of cross-sectional earnings. According to Table 8, the earnings Gini rises from 0.310 to 0.324. Given that the earnings Gini under the 2017 calibration is 0.389, the college premium counterfactual alone accounts for 18 percent of the total change in cross-sectional earnings inequality between 1979

[56]Table 2 (the first block) lists these parameters and their corresponding values.

[57]Appendix E shows the marginal contribution of the remaining model parameters to the optimal policy. Among all model primitives, household utility and representative firm technology have the most significant quantitative impact on the optimal policy.

[58]See Table 2 (the first block) for the parameter values used in each experiment.

Table 8: INEQUALITY MOMENTS IN COUNTERFACTUAL EXPERIMENTS.

	Gini(e)	**Gini(y)**	**Corr($\bar{e}, y$)**
Baseline (1979)	0.310	0.296	0.76
Counterfactuals (2017):			
– *Premium*	0.324	0.346	0.84
– *Risk*	0.353	0.357	0.77

Notes: $\mathrm{Gini}(e)$ refers to pre-tax earnings. $\mathrm{Gini}(y)$ refers to incomes before taxes and government transfers. $\mathrm{Corr}(\bar{e}, y)$ is the correlation between average lifetime earnings and pre-government incomes. The last two statistics are calculated based on the sample of agents who enter retirement (i.e., agents of age $j = J^R$). All numbers are computed based on the pre-reform steady state distribution in 2017 with $\boldsymbol{\alpha}_{1979}^{data}$. In each counterfactual, all parameters are fixed at their calibrated values for 1979, except for the following parameters that take on values from 2017: the age-efficiency profiles, $\{\zeta_{z,j}\}_{z,j=1}^{J^R-1}$ (*Premium*) and the variances of the persistent and temporary shocks, $\{\sigma^2_{\gamma,z}, \sigma^2_{\nu,z}\}_z$ (*Risk*).

and 2017. Since $\zeta_{z,j}$ permanently shifts wages of high-ability workers upward, the disparity in cross-sectional earnings propagates into an unequal distribution of incomes at retirement. One can see this by studying the inequality in market incomes among those agents who enter retirement (i.e., all agents of age J^R).[59] Recall that the retired agents' only source of income (apart from government transfers) is the return on their savings. Hence, the market income inequality measures the dispersion in the accumulated wealth among retired agents. According to the table, the Gini index for market income rises by 17 percent (from 0.296 to 0.346), indicating a larger disparity in the agents' ability to provide for their retirement.

Table 9 shows the distribution of ability among retired agents and confirms the simple hypothesis that these are the ability differences in wages that drive inequality at retirement. More specifically, the table reports the share of high-ability agents per each quartile of the market income distribution. Whereas the fraction of high-ability agents in the top quartile of the market income distribution was 51.0 percent in 1979, it rises to 71.8 percent solely due to the college premium shift. Simultaneously, the same share in the bottom quartile falls from 13.3 to 2.7 percent. Hence, the distribution of high-ability agents has become more concentrated at higher incomes, compared to 1979.

Naturally, the government will seek ways to counteract inequality by intensifying redistribution from income-rich toward income-poor retirees. However, the government can only

[59] Qualitatively, none of the arguments presented in this section change if I instead focus on the full sample of retired agents. However, studying the agents who are just entering retirement draws a clearer picture of the retired agents' economic well-being because all agents, regardless of their financial status, reduce their assets as they approach the terminal period J.

Table 9: SHARES OF HIGH-ABILITY AGENTS, PER INCOME QUARTILE

	Q1	Q2	Q3	Q4
Baseline (1979)	13.3	12.6	23.2	51.0
Counterfactuals (2017):				
– *Premium*	2.7	4.4	21.3	71.8
– *Risk*	17.9	17.6	14.8	49.7

Notes: The table displays the shares of retirees with average lifetime earnings above the economy-wide average (i.e., retirees with $\hat{e} > 1$), by quartiles of the pre-government income distribution. The statistics are calculated based on the sample of agents who enter retirement (i.e., agents of age $j = J^R$). All numbers are computed based on the pre-reform steady-state distribution in 2017 under the Social Security policy $\boldsymbol{\alpha}_{1979}^{data}$. In each counterfactual, all model parameters are fixed at their calibrated values for 1979, except for the following parameters that take on values from 2017: the age-efficiency profiles, $\{\zeta_{z,j}\}_{z,j=1}^{J^R-1}$ (*Premium*) and the variances of the persistent and temporary shocks, $\{\sigma^2_{\gamma,z}, \sigma^2_{\nu,z}\}_z$ (*Risk*).

employ the replacement rate schedule, which conditions the agent's pension benefit on this agent's average lifetime earnings. Hence, the question is whether the income-poor retirees are those agents who have also accumulated low lifetime earnings during their working careers. Table 8 confirms that this is indeed the case. The correlation between pre-government incomes and average lifetime-earnings among retired agents jumps from 0.76 to 0.84. Table 10 draws a more detailed picture. It displays the share of agents at age J^R whose average lifetime earnings are above the economy-wide average taxable earnings (i.e., those agents with $\hat{e} > 1$) per quartile of the market income distribution. In the baseline model under the 1979 calibration, the share of these agents in the first, second, and third quartile are 4.7, 25.4, and 66.0 percent, respectively. By contrast, these shares plummet by roughly four, three, and two times, respectively, in the college premium counterfactual. Hence, the individual's average lifetime earnings become a better indicator of their economic well-being in this experiment.

Since the individual's average lifetime earnings become more indicative of their economic status, the government can effectively use the pension system progressivity to target pension benefits at the income-poor retirees. This explains why α_2 rises by 18.3 percent in the college premium counterfactual (Table 7). As pension benefits become more precisely targeted at the disadvantaged groups, it would be inefficient to grant income-rich retirees high replacement rates. This explains why the average replacement rate, α_1, declines by 11 percent. This measure benefits the entire working-age population because it reduces the distortions associated with labor taxation.

Table 10: SHARES OF RETIREES WITH AVERAGE LIFETIME EARNINGS ABOVE ECONOMY-WIDE AVERAGE EARNINGS, PER INCOME QUARTILE.

	Q1	Q2	Q3	Q4
Baseline (1979)	4.7	25.4	66.0	97.1
Counterfactuals (2017):				
– *Premium*	1.2	9.2	38.9	90.1
– *Risk*	5.9	23.7	63.3	97.7

Notes: The table displays the shares of retirees with average lifetime earnings above the economy-wide average (i.e., retirees with $\hat{e} > 1$), by quartiles of the pre-government income distribution. The statistics are calculated based on the sample of agents who enter retirement (i.e., agents of age $j = J^R$). All numbers are computed based on the pre-reform steady-state distribution in 2017 under the Social Security policy $\boldsymbol{\alpha}^{data}_{1979}$. In each counterfactual, all model parameters are fixed at their calibrated values for 1979, except for the following parameters that take on values from 2017: the age-efficiency profiles, $\{\zeta_{z,j}\}_{z,j=1}^{J^R-1}$ (*Premium*) and the variances of the persistent and temporary shocks, $\{\sigma^2_{\gamma,z}, \sigma^2_{\nu,z}\}_z$ (*Risk*).

Idiosyncratic risk

In the next counterfactual, I adjust the variances of the idiosyncratic labor productivity risk, while keeping all other model parameters fixed at the 1979 levels. Based on my calibration results reported in Table 2, the persistent shock's variance doubles, whereas the variance of the temporary shock increases by 25 percent for a high-ability agent. The corresponding numbers for a low-ability agent are 45 percent and 8 percent, respectively.

Ex-post, i.e., after the idiosyncratic risk has been realized, the inequality in cross-sectional earnings rises substantially. According to Table 8, the earnings Gini jumps from 0.310 in 1979 to 0.353 in 2017. Quantitatively, this amounts to 54 percent of the total rise in inequality between 1979 and 2017. Similar to the college premium counterfactual, there is a substantial passthrough of inequality from the working stage into retirement. Indeed, the same table shows that the income Gini surges from 0.296 to 0.357.

However, there are some notable differences from the previous experiment. First, the distribution of high-ability agents by income shifts to the *left*, as shown in Table 9. More specifically, the shares of high-ability agents in the top two income quartiles fall, whereas the opposite happens at the bottom two quartiles. The most pronounced drop occurs at the third quartile, where the share of high-ability agents reduces from 23.2 to 14.8 percent. Second, contrary to the previous experiment, there is a weaker correlation between retired agents' incomes and their average lifetime earnings (0.77 compared to 0.84). Table 10 sheds more light on this observation by zooming onto the retired agents with average lifetime earnings above the economy-wide average earnings. The table shows that the distribution of these agents by income remains almost unchanged, compared to 1979.

The optimal policy shown in Table 7 reflects these two major differences from the college premium experiment. In response to more volatile labor productivity, high-ability workers can expect a higher probability of falling into the lower tail of the average lifetime earnings distribution at retirement. As their demand for publicly provided insurance increases, the government optimally chooses to increase progressivity (by 24.3 percent) compared to 1979. However, it has to uphold roughly the same average replacement rate to redistribute to those retirees who have been lucky to accumulate relatively high average lifetime earnings but failed to save a sufficient amount of wealth.

College graduates

In the final counterfactual, the only parameter updated is the share of high-ability agents, λ_H. Recall that it increased from 25 percent in 1979 to 43 percent in 2017. As opposed to the previous two counterfactuals, the mechanism in this experiment operates through the density function, f_{2017}, in the government's optimality condition (22). As previously established (Section 4.1), high-ability agents across all age groups prefer a less progressive system than in the data. Hence, as their population size increases, the government's optimal response should be to reduce progressivity to cater to these agents' preferences. Indeed, Table 7 shows that α_2 goes down by 5.3 percent compared to 1979. The government complements this choice by reducing the average replacement rate by 8.9 percent. As the same table reveals, this measure dampens the distortionary pressure of labor taxation by reducing the Social Security tax rate by 0.8 percentage points in the long run. This benefits all working-age agents, especially those with high-ability because their pre-tax wages are higher, on average.

Summarizing, the available instruments allow the government to respond effectively to the increased college premium, while this ability is limited in the case of idiosyncratic risk. Overall, the effect of college premium and idiosyncratic risk on the optimal policy dominates quantitatively. Hence, the government optimally chooses to reduce the average replacement rate and increase progressivity.

5. Outlook

The analysis in this paper is subject to several critique points. I mention two of them below. Addressing each of these points opens exciting new avenues for future research.

First, I identify the shift in government preferences over insurance and redistribution but abstain from exploiting this information in further analysis. There could be a benefit from doing so. As already mentioned in the introduction, a large strand of economic literature studies the macroeconomic and welfare consequences of different retirement financing reforms. It remains unclear, however, to what extent the discussed reforms are feasible from a political standpoint. The identified distribution of Pareto weights can be applied in policy analyses to restrict all economically feasible proposals to those that are also politically viable.

Second, the shift in government preferences is *one* explanation for why Social Security has not adjusted to rising earnings inequality. Admittedly, there are alternative justifications. One such explanation is political gridlock. It describes a situation in which politicians fail to reach an agreement during the post-election bargaining stage. As a consequence, the policy remains at status-quo. Piguillem and Riboni (2016) have an interesting application of this mechanism to capital taxation. Their paper's mechanism might provide micro-foundations for the persistence of particular policies, including Social Security, despite the change in fundamentals.

References

Bartels, L., 2009. Unequal Democracy: The Political Economy of the New Gilded Age. Princeton University Press.

Bell, F.C., Wade, A.H., Goss, S.C., 1992. Life Tables for the United States Social Security Area. Actuarial Study 116, 1900–2100.

Benabou, R., 2000. Unequal Societies: Income Distribution and the Social Contract. American Economic Review , 96–129.

Bound, J., Geronimus, A., Rodriguez, J., Waidmann, T., 2014. The Implications of Differential Trends in Mortality for Social Security Policy. Working Papers 314. University of Michigan, Michigan Retirement Research Center.

Brendler, P., 2020. Why Hasn't Social Security Changed since 1977? Review of Economic Dynamics 36, 134 – 157.

Carceles-Poveda, E., Abraham, A., Unpublished. Tax Reform with Endogenous Borrowing Limits and Incomplete Asset Markets.

Chang, B.H., Chang, Y., Kim, S.B., 2018. Pareto Weights in Practice: A Quantitative Analysis Across 32 OECD Countries. Review of Economic Dynamics 28, 181 – 204.

Chang, B.H., Chang, Y., Kim, S.B., 2021. Ex-Ante Heterogeneity and Equilibrium Redistribution. Working Paper.

Conesa, J.C., Krueger, D., 1999. Social Security Reform with Heterogeneous Agents. Review of Economic Dynamics 2, 757–795.

De Nardi, M., 2004. Wealth Inequality and Intergenerational Links. The Review of Economic Studies 71, 743–768.

Elo, I.T., Preston, S.H., 1996. Educational Differentials in Mortality: United States, 1979–1985. Social Science & Medicine 42, 47 – 57.

Fehr, H., Habermann, C., 2008. Risk Sharing and Efficiency Implications of Progressive Pension Arrangements. The Scandinavian Journal of Economics 110, 419–443.

Fehr, H., Kallweit, M., Kindermann, F., 2013. Should Pensions Be Progressive? European Economic Review 63, 94 – 116.

Fuster, L., Imrohoroglu, A., Imrohoroglu, S., 2007. Elimination of Social Security in a Dynastic Framework. The Review of Economic Studies 74, 113–145.

Guvenen, F., 2009. An Empirical Investigation of Labor Income Processes. Review of Economic Dynamics 12, 58–79.

Hansen, G.D., 1993. The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked. Journal of Applied Econometrics 8, 71–80.

Heathcote, J., Perri, F., Violante, G.L., 2010a. Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006. Review of Economic Dynamics 13, 15–51.

Heathcote, J., Storesletten, K., Violante, G.L., 2010b. The Macroeconomic Implications of Rising Wage Inequality in the United States. Journal of Political Economy 118, 681–722.

Heathcote, J., Storesletten, K., Violante, G.L., 2017. Optimal Tax Progressivity: An Analytical Framework. The Quarterly Journal of Economics 132, 1693–1754.

Heathcote, J., Storesletten, K., Violante, G.L., 2020. How Should Tax Progressivity Respond to Rising Income Inequality? Working Paper 28006. National Bureau of Economic Research.

Hosseini, R., Shourideh, A., 2019. Retirement financing: An optimal reform approach. Econometrica 87, 1205–1265.

Huggett, M., Parra, J.C., 2010. How Well Does the U.S. Social Insurance System Provide Social Insurance? Journal of Political Economy 118, 76–112.

Huggett, M., Ventura, G., 1999. On the Distributional Effects of Social Security Reform. Review of Economic Dynamics 2, 498–531.

Jacobs, B., Jongen, E.L., Zoutman, F.T., 2017. Revealed Social Preferences of Dutch Political Parties. Journal of Public Economics 156, 81–100.

Kitao, S., 2014. Sustainable Social Security: Four Options. Review of Economic Dynamics 17, 756–779.

Lindbeck, A., Weibull, J.W., 1987. Balanced-Budget Redistribution as the Outcome of Political Competition. Public Choice 52, 273–297.

Lockwood, B.B., Weinzierl, M., 2016. Positive and Normative Judgments Implicit in U.S. Tax Policy and the Costs of Unequal Growth and Recessions. Journal of Monetary Economics 77, 30–47.

McGrattan, E.R., Prescott, E.C., 2017. On Financing Retirement with an Aging Population. Quantitative Economics 8, 75–115.

Meara, E.R., Richards, S., Cutler, D.M., 2008. The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 1981–2000. Health Affairs 27, 350–360.

Mendoza, E.G., Razin, A., Tesar, L.L., 1994. Effective Tax Rates in Macroeconomics: Cross-country Estimates of Tax Rates on Factor Incomes and Consumption. Journal of Monetary Economics 34, 297 – 323.

Moser, C., Olea de Souza e Silva, P., 2019. Optimal Paternalistic Savings Policies. Working Paper 17-51.

Ndiaye, A., 2020. Flexible Retirement and Optimal Taxation. Federal Reserve Bank of Chicago Working Paper 2018-18.

Nishiyama, S., 2019. The Joint Labor Supply Decision of Married Couples and the U.S. Social Security Pension System. Review of Economic Dynamics 31, 277 – 304.

Nishiyama, S., Smetters, K., 2007. Does Social Security Privatization Produce Efficiency Gains? The Quarterly Journal of Economics 122, 1677–1719.

Piguillem, F., Riboni, A., 2016. Dynamic Bargaining over Redistribution in Legislatures. Working Papers 2016-15. Center for Research in Economics and Statistics.

Rosenstone, S., Hansen, J., 1993. Mobilization, Participation, and Democracy in America. Macmillan, New York.

Social Security Administration, 2019. Annual Statistical Supplement To The 2019 Annual Report Of The Board Of Trustees Of The Federal Old-Age And Survivors Insurance And Federal Disability Insurance Trust Fund. Technical Report. The Board Of Trustees.

Stantcheva, S., 2016. Comment on Positive and Normative Judgments Implicit in U.S. Tax Policy and the Costs of Unequal Growth and Recessions by Benjamin Lockwood and Matthew Weinzierl. Journal of Monetary Economics 77, 48–52.

Tauchen, G., 1986. Finite State Markov-chain Approximations to Univariate and Vector Autoregressions. Economic letters 20, 177–181.

Trabandt, M., Uhlig, H., 2011. The Laffer Curve Revisited. Journal of Monetary Economics 58, 305 – 327.

Tsujiyama, H., Heathcote, J., 2015. Optimal Income Taxation: Mirrlees Meets Ramsey. Staff Report 507. Federal Reserve Bank of Minneapolis.

Waldron, H., 2007. Trends in Mortality Differentials and Life Expectancy for Male Social Security Covered Workers, by Socioeconomic Status. Social Security Bulletin 67.

Wu, C., 2020. More Unequal Income but Less Progressive Taxation. Technical Report.

Appendix A. Competitive equilibrium

First, I define the competitive equilibrium with an exogenous policy. Then I define the equilibrium, in which Social Security arises endogenously.

Definition 1. Given initial policy, $\boldsymbol{\alpha}_1$, the initial stock of capital, K_1, the initial distribution of agents F_1, the pre-existing maximum taxable earnings thresholds $(cap_{SS,0}, cap_{M,0})$ and average taxable earnings $\tilde{E}_0$, a **competitive equilibrium with an exogenous policy** $\boldsymbol{\alpha}$ is a sequence of individual functions for households, $\{\tilde{V}_t, V_t, c^\star, l^\star, a'^\star, \{\bar{u}_{j-34,t}\}_{j=35}^{J}\}_{t=1}^{\infty}$, a sequence of production plans for firms, $\{K_t, N_t, Y_t\}_{t=1}^{\infty}$, total wealth held by households, $\{A_t\}_{t=1}^{\infty}$, prices, $\{w_t, r_t\}_{t=1}^{\infty}$, Social Security tax rate and deficit, $\{\tau_{SS,t}, D_t\}_{t=1}^{\infty}$, wasted government spending $\{G_t\}_{t=1}^{\infty}$, maximum taxable thresholds, $\{cap_{SS,t}, cap_{M,t}\}_{t=1}^{\infty}$, taxable earnings $\{\tilde{E}_t\}_{t=1}^{\infty}$, Medicare transfers, $\{T_{M,t}\}_{t=1}^{\infty}$, distributions of agents, $\{F_t\}_{t=2}^{\infty}$, such that the following statements hold for all t:

- Functions $(\tilde{V}_t, V_t, c^\star, l^\star, a'^\star, \{\bar{u}_{j-34,t}\}_{j=35}^{J})$ solve agent's optimization problem in (11), where the average utility of children at age $\bar{j} \geqslant 35$ reads:

$$\bar{u}_{\bar{j},t} = \frac{\int_{x:j=\bar{j}} u(c^\star, l^\star)\, dF_t(x)}{\int_{x:j=\bar{j}} dF_t(x)}. \tag{A.1}$$

- Factor prices r_t and w_t are determined competitively from (2);
- Social Security budget constraint is given by (15);
- Government stock of debt in Social Security as a fraction of GDP remains constant:

$$d_{SS} = D_t / Y_t, \tag{A.2}$$

  where aggregate output Y_t is given by (1).
- Government wasted spending follows from the government budget constraint (18).
- Aggregate consumption, aggregate stock of supplied assets and total bequeathed wealth

are defined as follows:

$$\begin{aligned} C_t &= \int_{\mathbf{x}} c^{\star} dF_t(\mathbf{x}), & \text{(A.3)} \\ A_t &= \int_{\mathbf{x}} a dF_t(\mathbf{x}), & \text{(A.4)} \\ \Phi_t &= \int_{\mathbf{x}} (1-\psi_{j,z}) a'^{\star} dF_t(\mathbf{x}). & \text{(A.5)} \end{aligned}$$

- Medicare runs a balanced budget according to (19).
- Average taxable earnings, $\tilde{E}_t$, are given by:

$$\tilde{E}_t = \int_{\mathbf{x}:j<J^R} \tilde{e}_{SS} dF_t(\mathbf{x}) \Big/ \mu^W, \tag{A.6}$$

- Maximum taxable earnings thresholds evolve according to:

$$\begin{aligned} cap_{SS,t+1} &= cap_{SS,t} \times \tilde{E}_{t+1}/\tilde{E}_t, & \text{(A.7)} \\ cap_{M,t+1} &= cap_{M,t} \times \tilde{E}_{t+1}/\tilde{E}_t. & \text{(A.8)} \end{aligned}$$

- Capital, labor and goods markets clear:

$$\begin{aligned} K_t &= A_t - D_t, & \text{(A.9)} \\ N_t &= \int_{\mathbf{x}:j<J^R} \epsilon l^{\star} dF_t(\mathbf{x}), & \text{(A.10)} \\ C_t + G_t + K_{t+1} &= Y_t + (1-\delta)K_t. & \text{(A.11)} \end{aligned}$$

- Law of motion for f_t for $j = 1, ..., J-1$ is given by:

$$f_{t+1}(\mathbf{x}') = f_{t+1}(z, j+1, \mathbf{y}', \tilde{a}', \bar{e}') = \frac{\psi_{j,z}}{1+n} \int_{\mathbf{x}:j} \mathbb{1}_{\tilde{a}'=a_t'^{\star}} \pi(\mathbf{y}' \mid \mathbf{y}) dF_t(\mathbf{x}) \tag{A.12}$$

together with the distribution for age 1 agents:[60]

$$f_t(z = H, 1, \mathbf{y}, 0) = \lambda_H \Pi_y$$

[60]Recall that agents enter the model at age $j = 1$ without any assets.

and

$$f_t(z = L, 1, \mathbf{y}, 0) = (1 - \lambda_H)\Pi_y$$

where λ_H is the measure of newborn agents with high-ability and Π_y is the stationary measure of newborn agents with productivity $\mathbf{y}$.

Definition 2. A **competitive equilibrium with an endogenous policy** is:

- the paths of variables

$$\{V_t, \tilde{V}_t, c^\star, l^\star, a'^\star, \{\bar{u}_{j-34,t}\}_{j=35}^{J}, K_t, N_t, Y_t, A_t, w_t, r_t, \tau_{SS,t}, D_t, G_t\}_{t=1}^{\infty},$$

$$\{cap_{SS,t}, cap_{M,t}, \tilde{\bar{E}}_t, T_{M,t}, F_t\}_{t=1}^{\infty}$$

  and a policy vector $(\boldsymbol{\alpha}_1, \bar{\boldsymbol{\alpha}})$, which satisfy the definition of a competitive equilibrium with an exogenous policy (definition 1);

- Pareto-weights ω, such that $\boldsymbol{\alpha} = \boldsymbol{\alpha}^\star$, where $\boldsymbol{\alpha}^\star$ is the solution to the government's problem in (16).

Appendix B. Computational algorithm

Appendix B.1. Steady state equilibrium

Choose the grid points for asset holdings (a) with the number of grid points $N_a = 200$. The asset holdings are in the range $[0, 22]$. Using Tauchen (1986), create an age-dependent grid of dimension $(J^R - 1) \times 12$ for the idiosyncratic productivity shock $(\mathbf{y})$ taking into account the age-dependent variance of the persistent shock $\eta_{i,j}$ in (6). Make a guess on the age- and ability dependent grid for the average lifetime earnings $(\bar{e})$. At each age and ability level, I use 5 grid points for $\bar{e}$.

Follow the steps below to compute a steady state equilibrium with an exogenous policy $\boldsymbol{\alpha}$.

1. Make initial guesses of the steady state values of aggregate capital stock K, aggregate effective labor N, lump-sum Medicare transfer T_M, Social Security tax rate τ_{SS}, taxable earnings $\tilde{\bar{E}}$ and a vector of average instantaneous utilities $\{\bar{u}_{s-34}\}_{s=35}^{J}$. Make a guess on the grid for average lifetime earnings $\bar{e}$.
2. Given the guesses on K and N, compute prices w, r from (2) and total output Y from (1). Given Y, compute the implied Social Security deficit D using the constant government debt-to-GDP ration from (A.2).

3. Compute household's optimal labor and savings functions $l^\star$ and $a'^\star$, respectively, starting at age J and proceeding backwards.
4. Initialize the time-invariant distribution of agents $F(\mathbf{x})$, given that agents enter the model with zero assets and zero average lifetime earnings. Iterate the distribution forward using (A.12).
5. Compute aggregate assets A from (A.4). Given A and the guess on D, compute K using (A.9). Furthermore, compute N using (A.10), T_M using (19), τ_{SS} from the Social Security budget constraint in (15), $\tilde{E}$ from (A.6) and $\{\bar{u}_{s-34}\}_{s=35}^{J}$ using (A.1). Construct a new grid for lifetime earnings, such that agents are equally distributed across the five grid points at each age and ability level.
6. If the calculated values from the previous step are close to the guesses made in step 1, we have found a steady state equilibrium. Otherwise, update the guesses and repeat the steps until convergence.
7. To complete the solution, compute aggregate consumption C from (A.3), total bequests Φ from (A.5), total net income taxes using (21) and wasted government spending G from (18).

Appendix B.2. Transitional dynamics

At time $t = 1$, the government makes an unanticipated announcement that the policy will change once-and-for all in the following period. The optimization problem of the government in (16) requires solving for transitional dynamics from the initial steady state associated with $\boldsymbol{\alpha}_t^{data}$ to a new steady state associated with $\boldsymbol{\alpha}$. To do so, proceed as follows.

1. Compute the initial steady state associated with $\boldsymbol{\alpha}_t^{data}$ following the steps described in Appendix B.1. Proceed similarly to compute the final steady state associated with the candidate policy $\boldsymbol{\alpha}$. Denote the initial steady state quantities with an lower bar, e.g. $\underline{K}, \underline{C}$, etc. Denote the new steady state with an upper bar, e.g. $\overline{K}, \overline{C}$.
2. Assume that the transition is completed during T periods. Since the computation of Pareto weights described in Appendix B.3 relies on computing numerical derivatives evaluated at $\boldsymbol{\alpha}_t^{data}$, $\boldsymbol{\alpha}$ would be typically in a local neighborhood of $\boldsymbol{\alpha}_t^{data}$, so that $T = 65$ should be sufficient. Of course, for policies located further away from the initial steady state, a larger T would be necessary.
3. Compute the transitional paths of the variables from the initial to the final steady state:
   (a) Guess the paths of $\{K_t, N_t, T_{M,t}, \tau_{SS,t}, \tilde{E}_t, \{u_{s-34,t}\}_{s=35}^{J}\}_{t=1}^{T-1}$ with $K_1 = \underline{K}$ and $\tilde{E}_0 = \underline{\tilde{E}}$.

(b) Given the guessed paths of K_t and N_t, compute the paths of w_t, r_t, Y_t, D_t. Given the guess on the paths of $\tilde{E}_t$, compute the paths of $cap_{SS,t}, cap_{M,t}$ using (A.7)-(A.8).

(c) Proceeding backwards from period $T-1$ to period 1, compute the path of optimal savings and labor choices in all transition periods. Note that the continuation value at $T-1$ is $V_T = \overline{V}$.

(d) Using $\underline{F}$ and the paths of the decision rules computed in the previous step, find the time-path of the distribution, $\{F_t\}_{t=2}^{T-1}$ by iterating forwards.

(e) Compute the paths of $A_t, K_t, N_t, T_{M,t}, \tau_{SS,t}, \tilde{E}_t$ and the path of instantaneous utilities $\{\bar{u}_{s-34,t}\}_{s=35}^{J}$.

(f) If the newly computed paths are sufficiently close to the guessed ones in each period, we have found the solution. Otherwise, update the guesses and return to step 3. Proceed until convergence.

(g) To complete the solution, compute the paths of C_t, Φ_t, G_t and the total net income taxes.

4. Once the sequence has converged, check whether T from step 2 is sufficient by increasing T and checking whether the equilibrium paths are affected.

Appendix B.3. Estimating Pareto weights

This section explains how to estimate a vector of Pareto weight parameters $\boldsymbol{\kappa}_t$, such that the optimal policy, $\boldsymbol{\alpha}^\star$, from (16) is equal to the actual policy, $\boldsymbol{\alpha}_t^{data}$.

Before proceeding to the algorithm, note that $\boldsymbol{\kappa}_t$ must satisfy the following two first-order conditions:

$$A \equiv \int_{\mathbf{x}} \omega(j, z; \boldsymbol{\kappa}_t) \frac{\partial V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}, F_t)}{\partial \alpha_1} dF_t(\mathbf{x}) \bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^{data}} = 0 \tag{B.1}$$

$$B \equiv \int_{\mathbf{x}} \omega(j, z; \boldsymbol{\kappa}_t) \frac{\partial V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}, F_t)}{\partial \alpha_2} dF_t(\mathbf{x}) \bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^{data}} = 0. \tag{B.2}$$

Also, note that we evaluate the first-order derivatives of the value function at the true policy $\boldsymbol{\alpha}_t^{data}$ because this is what the solution has to be, by construction. Importantly, the initial distribution of agents across states, $F_t(\mathbf{x})$, does not depend on the choice of $\boldsymbol{\alpha}$, which simplifies the analysis. Hence, once one obtains the two derivatives of the value function and the initial distribution of agents, the parameter vector $\boldsymbol{\kappa}_t$ can be computed by solving the system consisting of the two optimality conditions above.

Follow the steps below to compute $\boldsymbol{\kappa}_t$:

1. Solve the model for the initial steady state equilibrium and obtain the stationary distribution of agents, $F_t(\mathbf{x})$ (Appendix B.1).
2. Solve the model with transitional dynamics from the initial steady state with $\boldsymbol{\alpha}_t^{data}$ to a new steady state with $\boldsymbol{\alpha}^1 = (\alpha_{1,t}^{data} - h_1, \alpha_{2,t}^{data})$ (Appendix B.2). h_1 is a sufficiently small step interval. I use $h_1 = 0.01$. Store the resulting value function $V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}^1, F_t)$.
3. Repeat the step above to solve the model with $\boldsymbol{\alpha}^2 = (\alpha_{1,t}^{data} + h_1, \alpha_{2,t}^{data})$. Store the resulting value function $V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}^2, F_t)$.
4. Compute $\partial V_t / \partial \alpha_1$ numerically:

$$\left.\frac{\partial V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}, F_t)}{\partial \alpha_1}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_t^{data}} = \frac{V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \overline{\boldsymbol{\alpha}}^2, F_t) - V_t(\mathbf{x}; \boldsymbol{\alpha}_t^{data}, \boldsymbol{\alpha}^1, F_t)}{2h_1}.$$

5. Proceeding similarly, compute numerically $\partial V_t / \partial \alpha_2$. I use the step size $h_2 = 0.1$.
6. Write down the following objective function:
   (a) The function takes as arguments some arbitrary value of $\boldsymbol{\kappa}_t$, the derivatives $\partial V_t / \partial \alpha_1$ and $\partial V_t / \partial \alpha_2$ computed in steps 2 and 3, respectively, as well as the distribution of agents F_t computed in step 1.
   (b) Given $\boldsymbol{\kappa}_t$, the function first computes the set of Pareto weights $\omega(j, z; \boldsymbol{\kappa}_t)$ using (17). For convenience, I use the agent's real-life age, $j + 24$, instead of the model age j as an argument of the weighting function.
   (c) Given $\omega(j, z; \boldsymbol{\kappa}_t)$, the function then computes the terms A and B in (B.1)-(B.2).
   (d) The output of the function is $A^2 + B^2$.
7. Pass the function constructed in the previous step to a standardized function minimization routine (such as *fmincon* in Matlab) and let the routine optimize over $\boldsymbol{\kappa}_t$.

Appendix B.4. Computing the utilitarian policy

This section explains how to compute the utilitarian policy reported in Table 6.

At time t (1979 or 2017), the economy is assumed to be in a steady state associated with policy $\boldsymbol{\alpha}_{1979}^{data}$ and the respective calibration of model parameters. Facing the initial steady state distribution of agents, f_t, the government makes an unanticipated announcement that it will implement a new policy which becomes effective in the following period and remains constant forever. When deciding on the policy, the government solves the maximization problem in (16) assigning equal weights to all households. To find the optimal utilitarian policy at time t, follow the steps below:

1. Compute the initial steady state associated with $\boldsymbol{\alpha}_t^{data}$ following the steps described in Appendix B.1. Store the distribution of agents, F_t.
2. Create an initial grid for the average replacement rate $\alpha_1 \in \mathcal{R}_+$ and the pension system progressivity $\alpha_2 \in \mathcal{R}$. This grid should be iteratively refined during the procedure.
3. For each combination of (α_1, α_2) from the grid, compute transitional dynamics from the initial steady state associated with $\boldsymbol{\alpha}_t^{data}$ to the final steady state associated with policy (α_1, α_2) following the steps described in Appendix B.2. For each transition, store the agent's value function $V_t(\boldsymbol{x}; \boldsymbol{\alpha}_t^{data}, (\alpha_1, \alpha_2), F_t)$.
4. Compute the vector of the aggregate social welfare:

$$\int_{\boldsymbol{x}} V_t(\boldsymbol{x}; \boldsymbol{\alpha}_t^{data}, (\alpha_1, \alpha_2), F_t) dF_t(\boldsymbol{x}).$$

5. Select the policy that leads to the highest welfare in the vector computed in the previous step.
6. Make sure that the policy from the previous step does not lie on any of the bounds of the grid. Refine the grids if necessary and return to step 3 until a given level of accuracy is reached.

Appendix C. Calibration

Appendix C.1. Social Security

In the United States, an individual's pension benefit depends on their average monthly indexed earnings (AIME). The AIME is the average monthly earnings over the highest-earning 35 years of an individual's working career, indexed to the economy-wide taxable earnings when the worker reaches age 62. A statutory replacement rate schedule maps the worker's monthly pension benefit to the worker's AIME. The schedule comprises three brackets with a constant marginal replacement rate of 90 percent, 32 percent, and 15 percent in the lowest, intermediate, and highest bracket, respectively. The upper bound on AIME is naturally given by the maximum taxable earnings threshold since only earnings below this threshold flow into the AIME's computation.

Figure C.8 shows the implied schedule of average replacement rates that applies to all workers who turned 62 in 2017. When constructing the figure, I convert worker's monthly earnings and the bend points into their annual counterparts (multiplying each by 12) since one period in my model corresponds to one year. The first two vertical dashed lines in the

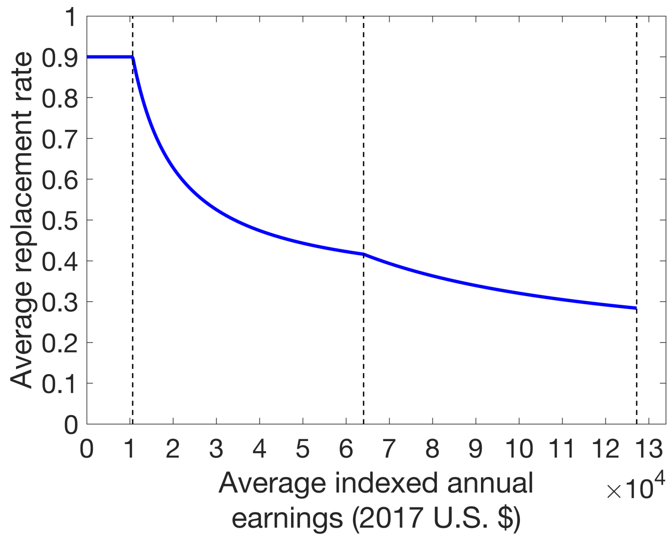

Figure C.8: Statutory Replacement Rate Schedule.

Notes: The figure visualizes the relationship between the individual's average indexed annual earnings and the statutory average replacement rate that applies to this individual. The figure is constructed based on the statutory pension benefit formula from 2017 (see Table 2.A11 in Social Security Administration (2019) for the formula and parameter values). The first two dashed lines in the figure separate the brackets. The statutory marginal replacement rates in the lowest, intermediate, and highest bracket are 90 percent, 32 percent, and 15 percent, respectively. The last dashed line corresponds to the maximum taxable earnings threshold from 2017 equal to $127,200.

figure separate the brackets. The third vertical line in the figure corresponds to the maximum taxable earnings threshold from 2017 equal to of \$127,200.

Since the adoption of the Social Security Amendments in 1977, the worker's AIME, the bracket limits, and the maximum taxable earnings threshold have been automatically indexed to the economy-wide average taxable earnings. Moreover, the statutory marginal replacement rates in each bracket have been constant since 1977. Consistent with these provisions, I adjust the worker's average lifetime earnings, $\bar{e}$, and the maximum taxable earnings threshold, cap_{SS}, in the model to the growth in economy-wide taxable earnings according to (8) and (A.7), respectively.[61]

I approximate the empirical replacement rate schedule using the specification from eq. (13). When doing so, I conduct three transformations on the empirical schedule. First, I convert worker's monthly earnings and the brackets into their annual counterparts (multiplying each by 12) since one period in my model corresponds to one year. Second, I normalize the average indexed annual earnings by the economy-wide average taxable earnings. The latter values can be easily computed from the CPS. This normalization avoids having to convert dollar amounts into model units.

The last transformation concerns the observation unit. Whereas the statutory schedule applies to *individuals*, the observation unit in my model is a *household*. I use household-level data instead of individual-level data to account for the insurance against idiosyncratic labor productivity risk among the household members. The literature showed this mechanism to be quantitatively important (see Fuster et al., 2007).[62] Hence, I adjust the brackets and the maximum taxable earnings threshold to account for the average number of earners in a household consistent with the CPS data. Below I explain the procedure for the 2017 calibration, but one can follow the same steps for 1979.

First of all, I compute the shares of households with more than one earner in the CPS. Consistent with the sample selection criteria applied in the paper, I define earners as those household members who work at least 260 hours per year. In 2017, there is less than 3 percent

[61] I deviate from the existing regulation in several ways. First, I compute earnings at an annual frequency. Second, I calculate average earnings over the agent's entire working career. Taking the average over the agent's 35 highest-earning years would be computationally infeasible. Finally, in the data, the worker's earnings are indexed to the economy-wide average taxable earnings in the year when the worker turns 62 years old. In the model, I index individual earnings to the period which precedes the worker's mandatory retirement.

[62] Without adjustment for household decomposition, the parameters of the statutory replacement rate schedule, $\boldsymbol{\alpha}$, would be constant during 1979–2017 due to the provisions stipulated by the Social Security Amendments of 1977. In this case, however, the paper's main results (e.g., the estimated trend in Pareto weights) would continue to hold qualitatively.

of households with four or more earners. I disregard these households in the calculations below. In the remaining sample, there is 47 percent of single-earner, 46 percent of two-earner, and 7 percent of three-earner households.

Second, I adjust the two bend points and the maximum taxable earnings threshold for the number of household earners. The bend points in the statutory replacement rate schedule are \$885 and \$5,336 (in terms of 2017 dollars).[63] I annualize these amounts to obtain \$10,620 and \$64,032, respectively. The maximum taxable earnings threshold is \$127,200. Adjusting the first bend point for the household composition, I obtain: 10,620 $\times$ [0.47 + 2$\times$0.46 + 3$\times$0.07] = \$16,992. I proceed similarly with the second bend point and the cap to get \$102,451 and \$203,520, respectively.

Next, I construct an updated schedule of average replacement rates based on the results obtained above. I apply the adjusted bend points and the cap and use the marginal replacement rates of 90, 32, and 15 percent in each bracket, respectively. Then I normalize the household's average annual lifetime earnings by the economy-wide average taxable earnings. In the CPS, the average household earnings capped at \$203,520 are \$87,837, which is the normalizing factor I use.

Finally, I fit the function in (13) to the constructed schedule of average replacement rates. The obtained estimates are $\boldsymbol{\alpha}_{2017}^{\mathtt{data}} = (0.5, 1.48)$, which are the values shown in Table2.

Given the calculations above, the maximum taxable earnings threshold (adjusted for the number of earners) was \$203,520, while the economy-wide average taxable earnings were \$87,837 in 2017. I calibrate $\mathtt{cap}_{SS,2017}$ inside the model to match the ratio of maximum taxable earnings threshold to the economy-wide average taxable earnings equal to 2.32. I proceed similarly for the 1979 calibration, where I find that this ratio to be 1.79.

Appendix C.2. Education-specific mortality differences

For the initial steady state in 1979, I assume that agents of both ability types face the same mortality rates at each age. I feed into the model the age-profile of conditional survival rates for Social Security workers estimated by Bell et al. (1992) for 1970. These data imply the age profile of life expectancy shown by the solid line in Figure C.9.

To calibrate $\psi_{z,j}$ for 2017, I rely on the data reported by Bound et al. (2014) (their Appendix 1). The authors use the data from the National Vital Statistics System and the Census from 1990–2010. They compute survival probability rates for a 25 year-old individual

[63] The parameters of the pension benefit formula can be found in Table 2.A11 in Social Security Administration (2019).

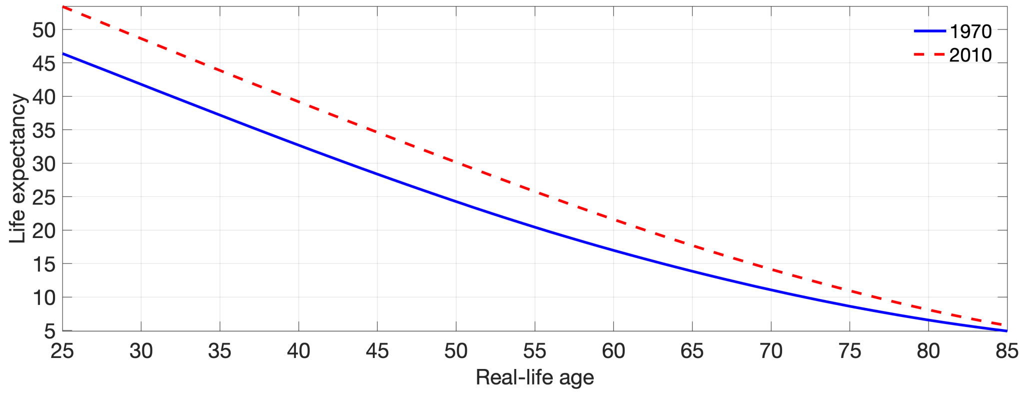

Figure C.9: Life Expectancy in the Data.

Notes: The figure displays the expected number of years of life remaining at a given real-life age for a Social Security covered worker. The data are computed by Bell et al. (1992).

for 1990, 2000 and 2010. I take their estimates for the most recent year 2010. Bound et al. (2014) report their estimates for 16 subsamples of individuals identified by gender, race, and education. With respect to race, they distinguish non-hispanic black and white individuals. Regarding education, they distinguish individuals with less than a high school degree, a completed high school degree, some years of college, and a complete college degree or higher.

Given these inputs, I proceed as follows. First of all, I pool together all individuals with an uncompleted college degree or lower. To compute the average profile of survival rates within this sample, I weight the probabilities of a given demographic group by its population share that I compute from the 2018 CPS extract. I proceed similarly to obtain the average profile of survival rates for the sample of individuals with a completed college degree or higher. The resulting empirical age profiles by education are shown in the lower panel of Figure C.10.

Since Bound et al. (2014) report survival probability rates for tabulated ages (25, 30, etc.), I fit a Gompertz force of mortality (i.e., mortality hazard) function to each of the two profiles of survival rates obtained above. This Gompertz function reads:

$$M_{zj} = \frac{\mu_0}{\vartheta_z} \times \frac{\exp(\mu_1 j)}{\mu_1}. \tag{C.1}$$

The Gompertz distribution is widely used in the actuarial literature and economics.[64] The second term in (C.1) controls how mortality changes with age, conditional on agent's ability. The first term determines the gradient of mortality in the cross-section. The key parameter

[64]See Hosseini and Shourideh (2019) and the references mentioned on p. 1224.

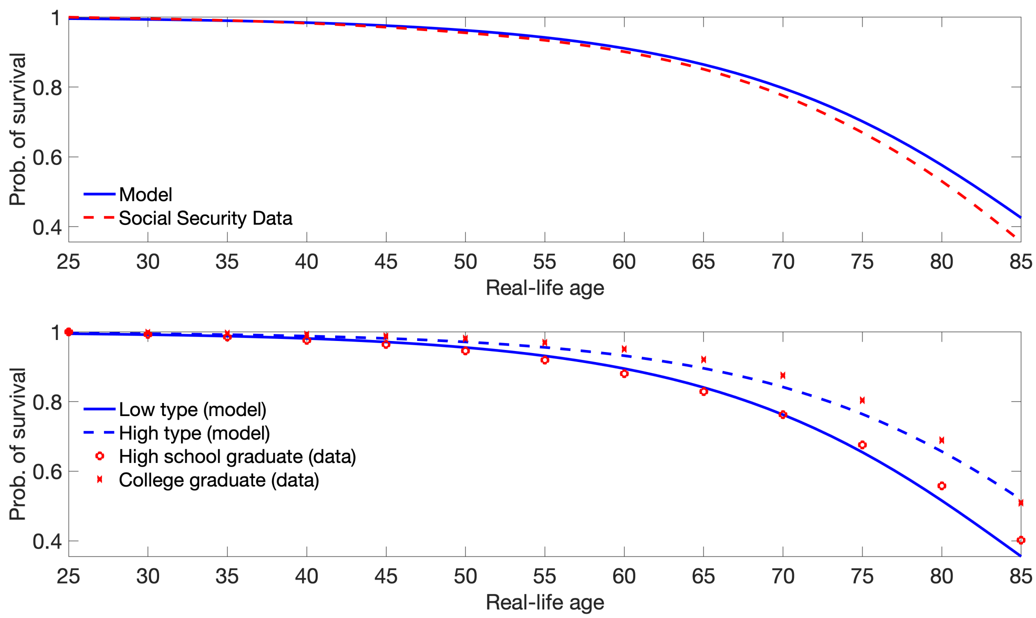

Figure C.10: Survival Probability Rates in the Model and Data.

Notes: Both panels compare the fit of survival probability rates. The top panel shows the rates for a 25-year old worker. The empirical rates (dashed line) were constructed based on the data by Bell et al. (1992) for 2010. These are the same data used to calculate empirical life expectancies in Figure C.9 (dashed line). The bottom panel compares the fit of survival rates for a 25 year-old worker of each education type in the model and data. The empirical moments (tabulated at 5-year intervals) are taken from Bound et al. (2014).

is ϑ_z which shows how mortality varies with agent's ability. Under this specification, the probability that an individual of age 1 survives up to age $j > 1$ is given by:

$$P_{zj} = \Pi_{i=1}^{j-1} \psi_{z,i} = \exp(-M_{z,j}). \tag{C.2}$$

Thus, the conditional survival probability rate, i.e., the variable ψ_{zj}, can be computed as: $\psi_{z,j} = P_{z,j+1}/P_{z,j}$ with $P_{z,1} \equiv 1$.

I normalize ϑ_H to one, so that the parameters to be estimated are $(\mu_0, \mu_1, \vartheta_L)$. I estimate these parameters to match the following targets. First, I target the two education-specific profiles of survival probability rates computed from Bound et al. (2014) at the tabulated ages (shown in the lower panel of Figure C.10). Second, I target the average profile of survival probability rates as of age 25 that I take from the period life table for the Social Security Area from Bell et al. (1992) for year 2010.[65] To compute the model counterpart of this profile, I use $M_j = \lambda_H M_{H,j} + (1-\lambda_H) M_{L,j}$, where λ_H is the cross-sectional shares of college graduates in the 2018 CPS extract equal to 43 percent. This profile is shown in the upper panel of Figure C.10.

The parameters that give the best fit are: $\mu_0 = 0.0003$, $\mu_1 = 0.0893$, $\vartheta_L = 0.6346$ and $\vartheta_H = 1.0$.

The same Figure C.10 shows the quality of the fit of the empirical targets. The top panel illustrates the fit of survival probability rates for a 25 year-old individual. The dashed line in the figure is constructed based on the same empirical estimates that I used to compute life expectancy in Figure C.9 (dashed line). The bottom panel shows the fit of education-specific profiles of survival rates computed for a 25 year-old individual.

Appendix D. Empirical evidence

This section provides empirical evidence in favor of the identified shift in Pareto weights toward older and high-ability agents (Table 6).

First of all, note that the social welfare function in (16) is equivalent to the micro-founded probabilistic voting environment introduced by Lindbeck and Weibull (1987). In this environment, two candidates maximize the probability of winning an election by proposing simultaneously and independently a (potentially, multidimensional) policy. Voters differ in their most preferred policies and other exogenous characteristics independent of the citizens'

[65] To compute this profile in the data, I weight equally the reported gender-specific profiles using as a weight the share of female household heads which is 44 percent, consistent with the CPS in 2017.

most preferred policies. The weight that both candidates attach to a given group is given by a product of the group's population size and its degree of homogeneity regarding exogenous characteristics. One such important characteristic is the group's turnout rate. The higher the group's propensity to vote, the higher the candidate's incentive to shift the policy closer to this group's most preferred policy. In equilibrium, both candidates propose the same policy.

In my model, the Pareto weights in the social welfare function can be interpreted as the weights that the candidates assign to electoral groups in the probabilistic voting environment. With this interpretation of the Pareto weights, the qualitative predictions of my model are as follows:

1. Conditional on age, the turnout rate of a college graduate relative to the high-school graduate's turnout must increase during 1979–2017.
2. Conditional on education, the turnout rate of older households relative to the young households' turnout must increase during 1979–2017.

To test these model predictions, I merge the CPS March Supplement with the survey data on voting behavior from the *Voting and Registration Supplement*, which is part of the CPS. It is most convenient to access the data through the *Integrated Public Use Microdata Series* (IPUMS) project webpage.[66]

In line with the sample selection criteria explained in the calibration section, a household head is a college graduate if she/he has a completed college degree. Otherwise, the household head is a high school graduate. I drop all households whose educational level is missing. Furthermore, I restrict the CPS sample to include only household heads aged 25–85, consistent with the model. Finally, I restrict the same to those households who answer *Yes* or *No* to the question: "*Have you voted in the most recent November election?*" (variable *VOTED*). Thus, I remove all those households from the sample who refuse to answer the question or claim they do not know the answer. When computing voter turnout rates, I weight observations using the variable *VOSUPPWT* from the *Voting and Registration Supplement.*

I report voter turnout statistics for Congressional elections because Congress would implement a Social Security reform. The first available Congressional election is from 1978, while the latest is from 2018. I split the data set into two subsamples: 1978–1986 and 2010–2018, each comprising three Congressional election cycles. In each subsample, I split households into four groups by household head's age (25–44, 45–54, 55–64, and 65–85) and two groups

[66]See https://cps.ipums.org/cps-action/data_requests/download. To assemble the raw data set, select the *Basic March CPS* data set and the *Voting and Registration Supplement.*

Table D.11: Turnout Rates of College Graduates Relative to High-School Graduates, by Age.

Age	1978–1986	2010–2018	% change
25–44	60.8	67.9	4.5
45–54	38.4	43.9	4.0
55–64	30.7	38.3	5.8
65–85	34.1	30.2	−2.9

Notes: The central two columns of the table show the turnout rates of college graduates normalized by the turnout rates of high school graduates, by age (in percent). The last column shows the percentage change in relative voter turnout rates between 1978–1986 and 2010–2018. The table is constructed based on the *CPS March Supplement* merged with the survey data on voting behavior from the *Voting and Registration Supplement*. The turnout rates are computed at the household level. Each subsample comprises three Congressional election cycles.

by household head's education.

To test the first model prediction, I proceed in three steps. First, I compute the average turnout rate for every age and education subgroup in each subsample period. For the sake of brevity, I do not report these results in the paper. Second, I ask: By how much percent does the computed turnout rate for college degree graduates exceeds that for high school graduates in a given age group in a given period? Table D.11 (second and third columns) reports the results. Finally, the last column shows the percentage change in the relative turnout rate across the two subsamples.

As one can see from Table D.11, college graduates vote at higher rates than high school graduates, since all numbers in the second and third columns are positive. This is true for all age groups and both subsample periods. Numerous empirical studies have already documented that participation among households in almost any form of political activity (including voting) rises with the households' level of education in the United States.[67] I add to this finding in the literature how relative turnout rates have changed over time. The last column of the table displays the percentage change in the relative turnout rate. When comparing the numbers in the second and third columns, one can see that the relative turnout rates increase for all age groups over time, except for the 65–85-year-olds. Overall, the empirical evidence largely supports the model's first qualitative prediction regarding the shift in Pareto weights toward high-ability agents, conditional on their age.

[67] See, among many others, Rosenstone and Hansen (1993), Benabou (2000), Bartels (2009).

Table D.12: TURNOUT RATES RELATIVE TO 25–44-YEAR-OLDS, BY AGE AND EDUCATION.

	1978–1986		2010–2018		% change	
	COL	**HS**	**COL**	**HS**	**COL**	**HS**
45–54	22.8	42.6	23.5	43.9	0.5	0.9
55–64	30.7	60.7	35.8	64.7	3.9	2.5
65–85	33.5	60.0	41.6	83.2	6.1	14.5

Notes: The table shows the normalized turnout rates of college graduates (*COL*) and high-school graduates (*HS*), by age. Within each education type and time period, the turnout rates are normalized by the turnout rate of 25–44-year-olds of the same education level. All numbers are in percent. The last two columns show the percentage change in relative voter turnout rates between 1978–1986 and 2010–2018. The table is constructed based on the *CPS March Supplement* merged with the survey data on voting behavior from the *Voting and Registration Supplement*. The turnout rates are computed at the household level. Each subsample comprises three Congressional election cycles.

To test the second model prediction, I start with the same data set constructed in the first step above. Next, I ask: Conditional on household's education, by how much percent does the turnout rate of older households differ from the youngest (25–44) group? Table D.12 reports the results (columns `Col` and `HS`) for each time frame. Finally, I calculate the percentage change in the relative turnout rates across time within each education sample.

According to Table D.12, all three age groups (45–54, 55–64, and 65–85) voted at higher rates than the youngest group (25–44). This observation holds at each education level and in each time frame. These facts are well-known in the literature (see the sources cited above). Next, observe that the relative turnout rates rise at a higher rate for older groups within each education type over time. This empirical evidence supports the model's second prediction regarding the shift in Pareto weights toward older agents, conditional on their ability.[68]

Appendix E. Optimal policy in counterfactual experiments

The current section decomposes the parameters' effect on the optimal policy in 2017. All policies in this section are computed under the assumption that the Pareto weights remain fixed at the 1979 level. I split all model parameters that take on different values between 1979 and 2017 into six subgroups listed in Table 2 on page 20: 1) Inequality, 2) Social Security,

[68] In my previous work (Brendler, 2020), I have documented the rise in relative turnout rates by age using a *pooled* sample of college and high-school graduates. In the current work, I show that the results also hold for each education group taken separately.

3) Utility and technology, 4) Aging, 5) Income taxes, and 6) Other taxes. In the paper's main part (Section 3), I have already analyzed the combined effect of all parameters on the optimal policy. Moreover, I have quantified the marginal impact of earnings inequality (i.e., the first subgroup) while keeping all the remaining parameters fixed at the 1979 levels. I have established that, quantitatively, the increased earnings inequality explains 38.5 percent of the total decline in the average replacement rate and 32.5 percent of the total increase in progressivity. To investigate what model parameters account for the remaining portion of the optimal policy change, I proceed as follows.

When analyzing the effects of Social Security parameters (i.e., the second subgroup), I *maintain* the inequality parameters at their 2017 levels and *additionally* update those parameters that control Social Security. All the remaining parameters remain constant at the calibrated 1979 levels. Similarly, when studying the effect of utility and technology parameters (i.e., the third subgroup), I keep the inequality *and* the income tax parameters at the 2017 levels while leaving all other parameters at the 1979 levels. I proceed similarly with the remaining subgroups. This procedure allows me to evaluate the combined effect of inequality and other model parameters considering potentially complex general equilibrium interactions.

Table E.13 shows the results. Note that the table displays the optimal policy as a percent deviation from the calibrated policy in 1979. For the reader's convenience, the second and third columns replicate the results from the paper's main part. The table generates a set of interesting findings whose detailed exploration would go beyond this paper's scope. Instead, I will briefly highlight one observation only.

The combined effect of all model parameters leads to a reduction in the average replacement rate level by 28.9 percent and increases progressivity by 184 percent. The table reveals that the Social Security parameters and the utility and technology parameters explain the major portion of the total effect. For example, augmenting inequality with the shift in the Social Security parameters brings down $\alpha_1^\star$ by 20 percent and pushes up $\alpha_2^\star$ by 107 percent, thus explaining roughly 70 percent of the change in $\alpha_1^\star$ and 60 percent of the change in $\alpha_2^\star$.

Table E.13: Optimal Policy in 2017 in Counterfactual Experiments.

	All Changes	**Inequality**	**Inequality *and***				
			Social Security	**Utility and Technology**	**Aging**	**Income Taxes**	**Other Taxes**
$\alpha_1^\star$, %	−28.9	−11.1	−20.0	−53.3	−1.1	+5.6	−13.3
$\alpha_2^\star$, %	+184.0	+59.8	+107.1	+160.4	+30.2	+36.1	+65.7

Notes: The table shows the optimal policy in 2017 under the Pareto weights from 1979 in counterfactual experiments. The table displays all policies in percent deviations from the baseline model calibrated for 1979. *All Changes* reports the results when all model parameters take on values from 2017. *Inequality* updates those parameters that are listed in the first block of Table 2 while leaving all the remaining parameters at their calibrated levels for 1979. In all the remaining experiments, I keep the inequality parameters at their 2017 levels and additionally update the parameters from a given subgroup while keeping other parameters at the 1979 levels. Table 2 shows the subgroups.